CHAPTER 1

BASIC CONCEPTS OF PROBABILISTIC RISK ASSESSMENT (PRA)

1.0 INTRODUCTION

This chapter describes what a probabilistic risk assessment (PRA) is and compares and contrasts it to the more familiar point estimate methods described in EPA's *Risk Assessment Guidance for Superfund (RAGS) Part A* (U.S. EPA, 1989a). A risk assessment performed using probabilistic methods is very similar in concept and approach to the traditional point estimate method, with the main difference being the methods used to incorporate variability and uncertainty into the risk estimate. A variety of modeling techniques can be used to characterize variability and uncertainty in risk. This guidance focuses on Monte Carlo analysis (MCA), which is one of the most common probabilistic methods that risk assessors will encounter. Basic concepts on how to use MCA to propagate variability and uncertainty in exposure through a risk model are presented. In addition, the general advantages and disadvantages of both point estimate and probabilistic approaches are outlined. Many of the concepts presented in this guidance are applicable to other probabilistic approaches to risk assessment.

At some sites, probabilistic analysis may provide a more complete and transparent characterization of the risks and uncertainties in risk estimates than would otherwise be possible with a point estimate approach. However, the decision to conduct a PRA should be made after careful consideration of a variety of factors; the main decision points are presented in this chapter. Developing or reviewing a PRA may involve additional time and resources, and a PRA is not necessary or desirable for every site. The tiered approach presented in this chapter highlights the important scientific and management decisions for determining if PRA is appropriate at a specific site. The decision to perform PRA is best made after the Risk Assessor and the Remedial Project Manager at the site determine whether the available information will support a PRA and whether a PRA will enhance decision making at the site. If a PRA is conducted, the assumptions and inputs to the probabilistic model should be sufficiently documented so that the results can be independently reproduced.

1.1 BACKGROUND

EPA uses risk assessment (NAS, 1983; 1994) for compliance with the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), as amended by the Superfund Amendments and Reauthorization Act of 1986 (SARA). Under CERCLA/SARA, EPA's Superfund Program is authorized to protect human health and the environment from current and potential threats posed by uncontrolled releases of hazardous substances, pollutants, or contaminants. The primary regulation issued by the Superfund Program is the National Oil and Hazardous Substances Pollution Contingency Plan (NCP) (U.S. EPA, 1990). The NCP calls for the identification and mitigation of environmental impacts at hazardous waste sites, and for the selection of remedial actions to protect human health and the environment. An important part of the NCP is the implementation of a Remedial Investigation and Feasibility Study (RI/FS), which is designed to support risk management decisions within the Superfund Program. A risk assessment is an integral part of the RI/FS, and is generally conducted at a site to determine the need for action and to ensure that a selected remedy will be protective. To promote consistency and science-based approaches, EPA's Superfund Program has developed guidance on human health and ecological risk assessment. Risk Assessment Guidance for Superfund (RAGS) Volume 1 Part A (U.S. EPA, 1989a, b) provides guidance on gathering and assessing human health risk

information. Similar guidance is available for implementing a standardized process for ecological risk assessment (U.S. EPA, 1997a; 1998a; 1999). RAGS Volume 1 Part D (U.S. EPA, 1998b) provides guidance on the planning, reporting, and review of Superfund risk assessments.

EPA previously issued guidance that addresses the use of quantitative uncertainty analysis in risk assessment. RAGS Volume 1 (U.S. EPA, 1989a) and the Exposure Assessment Guidelines (U.S. EPA, 1992a) emphasize the importance of assessing variability and uncertainty in risk estimates conducted in the Superfund Program. Guidance is also available for characterizing the 95% upper confidence limit (UCL) for the mean exposure concentration (U.S. EPA, 1992b; 1997b). At the regional level, EPA Regions 3 and 8 issued guidance on the appropriate use of probabilistic methods in risk assessment (U.S. EPA, 1994a, 1995a). The importance of adequately characterizing variability and uncertainty is addressed in the 1995 memorandum on Risk Characterization Policy and Guidance (U.S. EPA, 1995b). In the spring of 1997, EPA released the memorandum, *Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment* (U.S. EPA, 1997c). According to the Policy Statement of the memorandum, probabilistic analysis techniques, "given adequate supporting data and credible assumptions, can be viable statistical tools for analyzing variability and uncertainty in risk assessments." As such, a PRA, "will be evaluated and utilized in a manner that is consistent with other risk assessments submitted to the Agency." Along with this Policy Statement, the Agency released a set of guiding principles for use and review of probabilistic analyses (U.S. EPA, 1997c). Hence, both RAGS and Agency-wide guidance emphasize the importance of review of the scientific and technical merit of a probabilistic analysis to determine whether or not the assessment is of sufficient quality to support a remedial decision. This guidance, RAGS Vol. 3, provides risk assessors with comprehensive guidance on when and how to conduct PRAs using Monte Carlo analysis within the Superfund Program.

1.2 KEY PRA TERMS - VARIABILITY AND UNCERTAINTY

An essential concept in PRA is the distinction between "variability" and "uncertainty". *Variability* refers to true heterogeneity or diversity. For example, among a population that drinks water from the same source and with the same contaminant concentration, the risks from consuming the water may vary. This may be due to differences in exposure (i.e., different people drinking different amounts of water, having different body weights, different exposure frequencies, and different exposure durations) as well as differences in response (e.g., genetic differences in resistance to a chemical dose). These inherent differences are referred to as ***variability***. Differences among individuals in a population are referred to as inter-individual variability, while differences for one individual over time is referred to as intra-individual variability.

Uncertainty occurs because of a lack of knowledge. It is not the same as variability. For example, a risk assessor may be very certain that different people drink different amounts of water, but may be uncertain about how much variability there is in water intakes within the population. Uncertainty can often be reduced by collecting more and better data, while variability is an inherent property of the population being evaluated. Variability can be better characterized with more data, but it cannot be reduced or eliminated. Efforts to clearly distinguish between variability and uncertainty are important for both risk assessment and risk communication.

1

EXHIBIT 1-1

DEFINITIONS FOR CHAPTER 1 (part 1 of 2)

Central Tendency Exposure (CTE) - A risk descriptor representing the average or typical individual in a population, usually considered to be the mean or median of the distribution.

Cumulative Distribution Function (CDF) - Obtained by integrating the PDF, gives the cumulative probability of occurrence for a random independent variable. Each value *c* of the function is the probability that a random observation *x* will be less than or equal to *c*.

Expected Value of Information (EVOI) - The expected increase in the value (or decrease in the loss) associated with obtaining more information about quantities relevant to the decision process. EVOI is a measure of the importance of uncertainty in risk and the potential for changing a risk management decision if uncertainty is reduced (see Appendix E).

Frequency Distribution or Histogram - A graphic (plot) summarizing the frequency of the values observed or measured from a population. It conveys the range of values and the count (or proportion of the sample) that was observed across that range.

Monte Carlo Analysis (MCA) or Monte Carlo Simulation - A technique for repeatedly sampling from probability distributions to derive a distribution of outcomes (e.g., risks)

Parameter - In PRA, a parameter is a constant that characterizes the probability distribution of a random variable. For example, a normal probability distribution may be defined by two parameters (e.g., arithmetic mean and standard deviation).

Point Estimate - In statistical theory, a quantity calculated from values in a sample to estimate a fixed but unknown population parameter. Such point estimates typically represent a descriptive statistic (e.g., arithmetic mean, 95^{th} percentile).

Point Estimate Risk Assessment - A risk assessment in which a point estimate of risk is calculated from a set of point estimates for exposure and toxicity. Such point estimates of risk can reflect, the CTE, RME, or bounding risk estimate depending on the choice of inputs.

Probability Density Function (PDF) - A function or graph representing the probability distribution of a continuous random variable. The density at a point refers to the probability that the variable will have a value in a narrow range about that point. Probability mass function refers to the probability distribution for a discrete random variable.

Probability Distribution - A mathematical representation of the function that relates probabilities with specified intervals of values for a random variable. Also called a *probability model*.

Probabilistic Risk Assessment (PRA) - A risk assessment that yields a probability distribution for risk, generally by assigning a probability distribution to represent variability or uncertainty in one or more inputs to the risk equation.

Uncertainty - Lack of knowledge about specific variables, parameters, models, or other factors (e.g., uncertainty regarding the concentration of a contaminant in an environmental medium, local fish consumption practices). Uncertainty may be reduced through further study.

1

EXHIBIT 1-1

DEFINITIONS FOR CHAPTER 1 (part 2 of 2)

Random Variable -A variable that may assume any of a set of values according to chance. Discrete random variables assume any value within a finite set of values, whereas continuous random variables can assume value within an interval.

Reasonable Maximum Exposure (RME) - The highest exposure that is reasonably expected to occur at a site (U.S. EPA, 1989a). The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.

Stochastic Dominance - Implies no intersection between the CDFs; distribution A stochastically dominates distribution B if, for every percentile of the CDF, A > B. This characteristic may not be apparent from the PDFs of the distributions, which may overlap.

Variability - True heterogeneity or diversity that characterizes an exposure variable or response in a population. Further study (e.g., increasing sample size, n) will not reduce variability, but it can provide greater confidence in quantitative characterizations of variability.

Toxicity Reference Value (TRV) - A risk-based dose or concentration that usually includes factors of toxic uncertainties and is often based on a NOAEL or LOAEL; a TRV is sometimes referred to as a "toxicity benchmark", but this is not the same as the conventional "benchmark dose", which is the lower probability bound on a dose for a designated low response.

2

1.3 WHAT IS PRA?

Probabilistic risk assessment (PRA) is a general term for risk assessments that use probability models to represent the likelihood of different risk levels in a population (i.e., variability) or to characterize uncertainty in risk estimates. In human health risk assessments, probability distributions for risk reflect variability or uncertainty in exposure. In ecological risk assessments, risk distributions may reflect variability or uncertainty in exposure or toxicity. A PRA that evaluates variability can be used to address the question, "What is the likelihood (i.e., probability) that risks to an exposed individual will exceed a regulatory level of concern?". For example, based on the best available information regarding exposure and toxicity, a risk assessor might conclude, "It is estimated that there is a 10% probability that an individual exposed under these circumstances has a risk exceeding $1x10^{-6}$." If a probabilistic approach also quantifies uncertainty, the output from a PRA can provide a quantitative measure of the confidence in the risk estimate. For example, a risk assessor might conclude, "While the best estimate is that there is a 10% chance that risk exceeds $1x10^{-6}$, I am reasonably certain (95% sure) that the chance is no greater than 20%."

A risk assessment performed using probabilistic methods relies on the same fundamental concepts and equations as the traditional point estimate approaches. *RAGS Volume 1: Part A* (U.S. EPA, 1989a) and the *Standard Default Factors Guidance* (U.S. EPA, 1991a) provide current guidance for estimating risk using the following standardized exposure and risk models (in Exhibit 1-2, units refer to exposure via drinking water):

EXHIBIT 1-2

CANCER AND NONCANCER RISK MODELS

Exposure Model:
$$CDI = \frac{C \cdot CR \cdot EF \cdot ED}{BW \cdot CSF}$$

Cancer Risk Model:
$$Risk = CDI \cdot CSF$$

Noncancer Risk Model:
$$HQ = \frac{CDI}{RfD}$$

where,

CDI = chronic daily intake of the chemical (mg/kg-day)
C = concentration of the chemical in an exposure medium (e.g., mg/L)
CR = contact rate (e.g. L/day for water ingestion, mg day^{-1} for incidental soil ingestion, etc.)
EF = exposure frequency (days/year)
ED = exposure duration (years)
BW = body weight (kg)
AT = averaging time (equal to ED x 365 days/year for noncarcinogens and 70 years x 365 days/year for carcinogens)
CSF = cancer slope factor (linear low-dose cancer potency factor) for the chemical (mg/kg-day)$^{-1}$
RfD = reference dose for the chemical for assessing non-cancer health effects (mg/kg-day)

In the point estimate approach, a single numerical value (i.e., point estimate) is chosen for each variable. For example, point estimates may include a drinking water ingestion rate of 2 L/day and a body weight of 70 kg for an adult. Based on the choices that are made for each individual variable, a single estimate of risk is calculated. In the probabilistic approach, inputs to the risk equation are described as *random variables* (e.g., variables can assume different values for different people) that can be defined mathematically by a probability distribution. For continuous random variables, such as those in Figure 1-1 (e.g., body weight), the distribution may be described by a probability density function (PDF), whereas for discrete random variables (e.g., number of fish meals per month), the distribution may be described by a probability mass function (PMF). The key feature of a PDFs and PMFs is that they describe the range of values that a variable may assume, and indicate the relative likelihood (i.e., probability) of each value. For example, drinking water ingestion might be characterized by a normal distribution with a mean of 2 L/day and a standard deviation of 1 L/day. After determining appropriate PDF types and parameter values for selected variables, the set of PDFs are combined with the toxicity value in the exposure and risk equations given above to estimate a *distribution* of risks. Chapter 3 provides guidance on selecting and fitting distributions.

At this time, for human health risk assessments, toxicity values will generally be characterized by point estimates because of limitations in the data and techniques for characterizing distributions for toxicity in humans. Only if adequate supporting data are available to characterize variability or uncertainty in toxicity values will the Agency consider the use of distributions for toxicity. The Agency will determine the adequacy of supporting data on a case-by-case basis, pending consultation with EPA Headquarters (Office of Emergency and Remedial Response). For ecological risk assessment, toxicity values may be characterized by probability distributions.

When displaying a continuous probability distribution in a graph, generally both the PDF and the corresponding cumulative distribution function (CDF) should be presented. To be most clear, it is recommended that a PDF and CDF be presented in adjacent (rather than overlaid) plots. Figure 1-1 illustrates a PDF and CDF for a normal probability distribution for adult body weight. Both curves represent the same distribution, but are useful for conveying different information. The CDF for risk can be especially informative for illustrating the percentile corresponding to a particular risk level of concern (e.g., $1x10^{-6}$). A text box may also be included on the graph to highlight important summary statistics, such as the parameters of the input distribution (e.g., Figure 1-1, 4-1), or selected percentiles of the output distribution for risk (e.g., Figure 4-2, 8-1). Important information for reporting probability distributions is summarized in Exhibit 1-3.

EXHIBIT 1-3

USE A PDF TO DISPLAY:

- the relative probability of values
- the most likely values (e.g., modes)
- the shape of the distribution (e.g., skewness, kurtosis, multimodality)
- small changes in probability density

USE A CDF TO DISPLAY:

- percentiles, including the median
- high-end risk range (e.g., 90th to 99th percentiles)
- confidence intervals for selected percentiles
- stochastic dominance (i.e., for any percentile, the value for one variable exceeds that of any other variable)
- mixed, continuous, and discrete distributions

*Source: EPA, 1997c

For example, a clear description of the parameters for the probability distribution should generally be given, as well as an indication of whether the distribution represents variability or uncertainty.

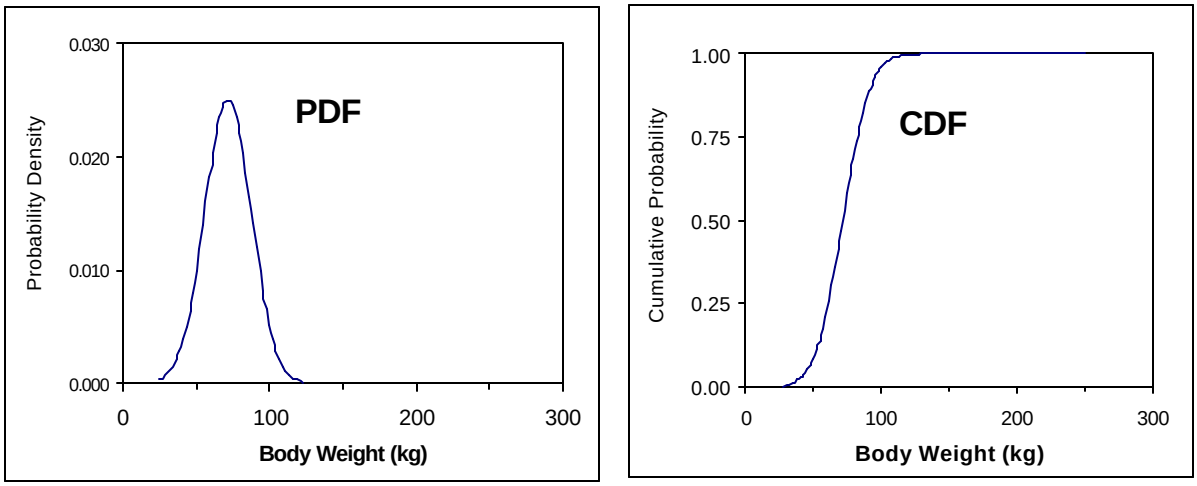


Figure 1-1. Example of a normal distribution that characterizes variability in adult body weight (males and females combined). Arithmetic mean = 71.7 kg, standard deviation = 15.9 kg (Finley et al., 1994). Body weight may be considered a continuous random variable. The left panel shows a bell-shaped curve and represents the probability density function (PDF), while the right panel shows an S-shaped curve and represents the cumulative distribution function (CDF).

In MCA (see next section) and related approaches to PRA, estimates for the probability distributions of assessment variables are combined to generate estimates of risk.

1.3.1 WHAT IS A MONTE CARLO SIMULATION?

The most common numerical technique for PRA is Monte Carlo simulation. The process for a Monte Carlo simulation is illustrated in Figure 1-2. In its general form, the risk equation can be expressed as a function of exposure and toxicity variables (V_i): Risk = f(V_1, V_2, …V_n). Solutions for equations with PDFs are typically too complex for even an expert mathematician to calculate the risk distribution analytically. However, computers can provide reasonably close approximations of a risk distribution using numerical techniques. This is illustrated here for the simplified case in which the assessment variables are statistically independent. In this case, the computer selects a value for each V_i at random from a specified PDF and calculates the corresponding risk. This process is repeated many times (e.g., 5000), each time saving the set of input values and corresponding estimate of risk. For example, the first risk estimate might represent a hypothetical individual who drinks 2 L/day of water and weighs 65 kg, the second estimate might represent someone who drinks 1 L/day and weighs 72 kg, and so forth. Each calculation is referred to as an iteration, and the set of iterations is called a simulation.

Each iteration of a Monte Carlo analysis represents a plausible combination of exposure and toxicity variables. A convenient aid to understanding the Monte Carlo approach for quantifying variability is to visualize each iteration as representing a single individual and the collection of all iterations as representing a population. In general, each iteration of a simulation should represent a plausible combination of input values, which may require using bounded or truncated probability distributions (see Chapter 3). However, risk estimates are not intended to correspond to any one person. The "individuals" represented by Monte Carlo iterations are virtual and the risk distributions derived from PRA allow for inferences to be made about the likelihood or probability of risks occurring within a specified range for an exposed human or ecological population. A simulation yields a set of risk estimates that can be summarized with selected statistics (e.g., arithmetic mean, percentiles) and displayed graphically using the PDF and CDF for the estimated risk distribution. More complex Monte Carlo simulations can be developed that quantify the correlations between one or more input distributions.

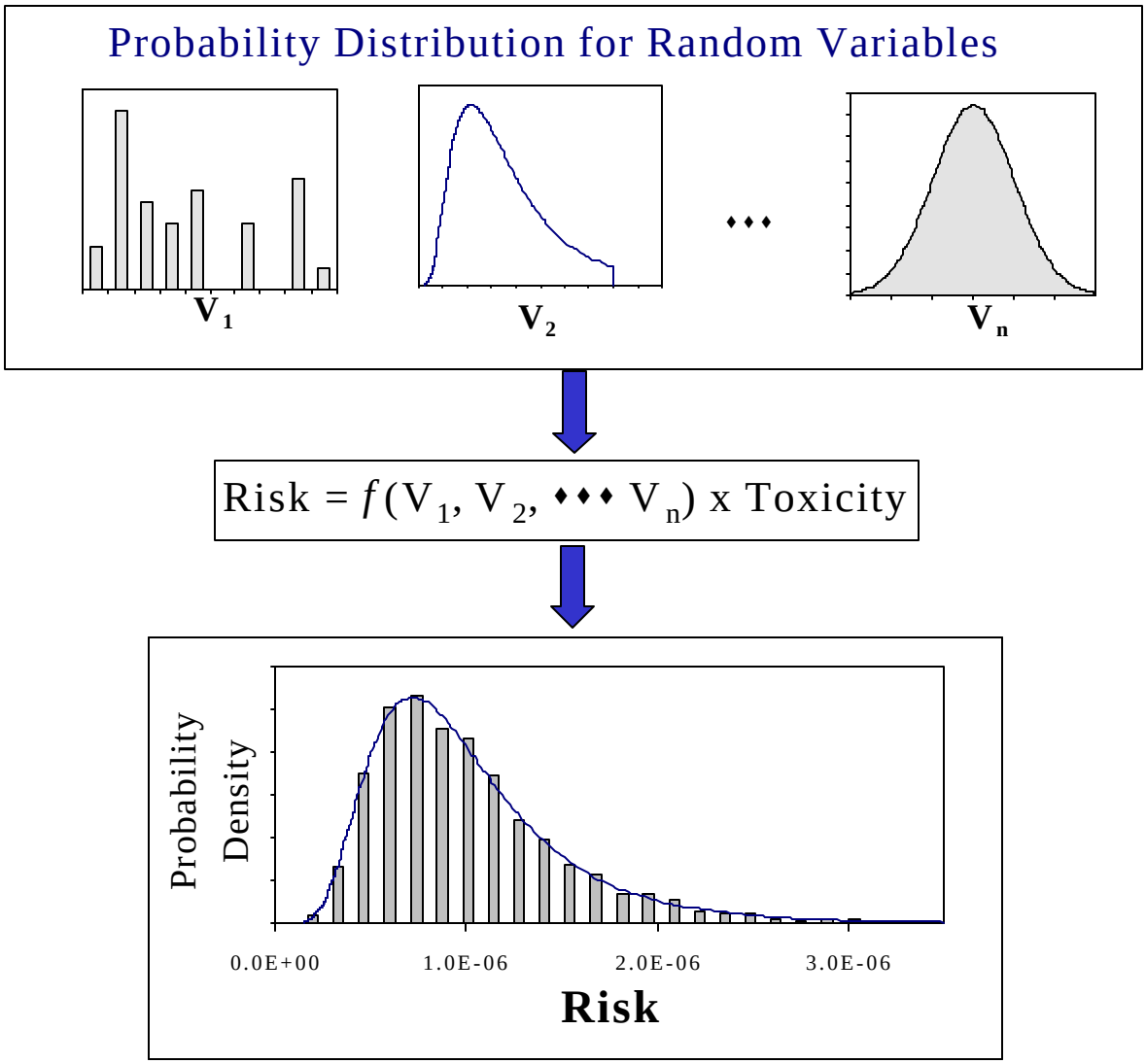


Figure 1-2. Conceptual model of Monte Carlo analysis. Random variables (V_1, V_2,…V_n) refer to exposure variables (e.g., body weight, exposure frequency, ingestion rate) that are characterized by probability distributions. A unique risk estimate is calculated for each set of random values. Repeatedly sampling $\{V_i\}$ results in a frequency distribution of risk, which can be described by a probability density function (PDF). The toxicity term should be expressed as a point estimate for human health risk assessment, but may be expressed by a probability distribution for ecological risk assessment.

1.3.2 WHY IS VARIABILITY IMPORTANT IN RISK ASSESSMENT? HOW IS IT ADDRESSED BY THE POINT ESTIMATE AND PROBABILISTIC APPROACHES?

As discussed previously, variability refers to true heterogeneity or diversity that occurs within a population or sample. For regulatory risk assessments, this variability can be characterized by a range of contaminant concentrations in a medium (air, water, soil, etc.), differences in intake rates or exposure frequencies, or in the case of ecological assessments, inter- and intra-species variability in dose-response relationships. Risk Assessment Guidance for Superfund (Section 6.1.2 of U.S. EPA, 1989a) and the NCP Preamble (U.S. EPA, 1990) state that Reasonable Maximum Exposure (RME) will generally be the principal basis for evaluating potential risks at Superfund sites. The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures based on both quantitative information and professional judgment (Sections 6.1.2 and 6.4.1 of U.S. EPA, 1989a). In addition, the Agency released guidance in 1992 (U.S. EPA, 1992c) recommending the inclusion of a "central tendency" exposure estimate to an individual, as well as a high-end exposure estimate, in the risk assessment. Generally, CTE is considered to be a measure of the mean or median exposure. For ecological risk assessments in which the endpoint of concern is local population effects (e.g., sustainability or community integrity) rather than individual-level effects, estimates of average or central tendency exposure in the population are typically desired (see Section 5.1.2).

Depending on assessment needs at a site, a range of point estimates of risk can be developed to represent variability in exposures. To support the evaluation of RME risk estimates using the point estimate approach described in Section 1.3, the Superfund Program developed guidance with recommended default values for exposure variables as inputs to the risk equations (U.S. EPA, 1992a). These standardized values are a combination of average (e.g., body weight, skin surface area) and high-end exposure assumptions (e.g., drinking water intake, exposure duration). A CTE risk estimate is based on central estimates (e.g., mean, 50^{th} percentile) for each of the exposure variables. Available site-specific data on plausible mean and upper range values for exposure variables should be used to support CTE and RME estimates. The point estimate approach does not allow the assessor to identify where the CTE or RME estimates lie among the risk distribution. For example, the RME based on the point estimate approach could be the 90^{th} percentile, the 99.9^{th} percentile, or some other percentile of the risk distribution. Without knowing what percentile is represented by the RME, the risk manager might be unsure about the level of protection corresponding to the RME.

In a PRA, distributions used as inputs to the risk equations may represent the inter-individual variability inherent in each of the exposure assumptions. By characterizing variability with one or more input distributions, the output from the Monte Carlo simulation is a distribution of risks, which represents the variability of exposures that could occur in that population. The high end of that risk distribution (e.g., 90^{th} - 99.9^{th} percentiles) is representative of exposures to the RME individual. The distribution mean and/or the central percentiles (e.g., 50^{th} percentile) can be associated with CTE risk estimates. In addition to providing a better understanding of where the RME and CTE risks occur in the distribution, a PRA also provides an estimate of the probability of occurrence associated with the risk (e.g., there is a 10% probability that risks exceed $1x10^{-6}$).

Provided that the arithmetic mean for each PDF used in a probabilistic approach is approximately equal to the point estimate for the corresponding variable, both approaches will likely yield similar estimates for the arithmetic mean (CTE) risk. Also, the point estimate calculation of the RME risk can be expected to lie somewhere in the upper end of the risk distribution, often between the 90^{th} and 99.9^{th} percentiles. The EPA Guidelines for Exposure Assessment (U.S. EPA, 1992a) states that the "high

end"of exposure for a population occurs between the 90th and 99.9th percentiles, with the 99.9th percentile considered a bounding estimate. While the correspondence between the results of two different approaches may provide a level of confidence in the risk assessment, care should generally be taken not to rely on such comparisons as a yardstick of the "accuracy" of either approach. If results of PRA calculations differ substantially from point estimate calculations, a risk manager may benefit from understanding the reasons for the differences and the relative strengths of the different approaches. For example, the point estimate approach may employ arithmetic mean input values for variables that are represented by extremely skewed distribution in the probabilistic approach. Sometimes, a closer look at the underlying assumptions and uncertainties in the different approaches will lead a risk assessor to revisit certain assumptions in order to provide a more consistent basis for comparison.

1.3.3 WHY IS UNCERTAINTY IMPORTANT IN RISK ASSESSMENT? HOW IS UNCERTAINTY ADDRESSED BY THE POINT ESTIMATE AND PROBABILISTIC APPROACHES?

Uncertainty is described as a lack of knowledge. The *Exposure Assessment Guidelines* (U.S. EPA, 1992a) and *Exposure Factors Handbook* (U.S. EPA, 1997d, e, f) describe a variety of different types of uncertainty in risk assessment as well as modeling strategies for addressing uncertainties. Potential sources of uncertainty are basically divided into those associated with model structure (e.g., mathematical equation) and those associated with assigning values to the parameters of the model. Some examples of the models that EPA uses in the risk assessment process are the equations to calculate exposure and risk, the linearized multistage model to assess cancer dose-response relationships, and media-specific models to estimate contaminant concentration. All models are simplified, idealized representations of complicated physical processes. They can be very useful from a regulatory standpoint, as it is generally not possible to adequately monitor long term exposure for populations at contaminated sites. However, models that are too simplified may not adequately represent all aspects of the phenomena they were intended to approximate or may not capture important relationships among input variables. Other sources of model uncertainty can occur when important variables are excluded, interactions between inputs are ignored, or surrogate variables that are different from the variable under study are used.

Parameter uncertainty may be the most readily recognized source of uncertainty that is quantified in site-specific risk assessments at hazardous waste sites. Parameter uncertainty can occur in each step of the risk assessment process from data collection and evaluation, to the assessment of exposure and toxicity. Sources of parameter uncertainty may include systematic errors or bias in the data collection process, imprecision in the analytical measurements, inferences made from a limited database when that database may or may not be representative of the variable under study, and

EXHIBIT 1-4

QUANTIFYING VARIABILITY AND UNCERTAINTY SIMULTANEOUSLY

1. **Single source of uncertainty.**
   Run multiple one-dimensional Monte Carlo simulations (1-D MCA), with each simulation using a different point estimate from the distribution for uncertainty. For example, uncertainty in the arithmetic mean concentration may be characterized by running separate simulations with the 95% lower confidence limit (LCL), sample mean, and 95% UCL.

2. **Multiple sources of uncertainty.**
   Run a single two-dimensional Monte Carlo simulation (2-D MCA), in which separate probability distributions are specified for variability and parameter uncertainty. See Appendices D and E for details.

extrapolation or the use of surrogate measures to represent the parameter of interest.

In the point estimate approach, parameter uncertainty is addressed in a qualitative manner for most variables. For example, the Uncertainty Section of a point estimate risk assessment document might state that a soil sampling plan yielded a small sample size that may not be representative of overall contaminant concentrations and, as a result, the risk estimate may over- or under-estimate actual risk. Uncertainty in the concentration term is addressed quantitatively to a limited extent in a point estimate approach. The Superfund program recommends using the 95% UCL for the arithmetic mean concentration in both CTE and RME risk estimates in order to account for uncertainty associated with environmental sampling and site characterization (U.S. EPA, 1992b; 1997b). The 95% UCL is combined in the same risk calculation with various central tendency and high-end point estimates for other exposure factors.

In a probabilistic approach, a probability distribution for risk will represent either variability or uncertainty, depending on how the distributions for the input variables are characterized. If exposure variability is characterized using probability distributions, the risk distribution represents variability. If input distributions represent uncertainty in estimates of central tendency (e.g., arithmetic mean), the output distribution represents uncertainty in the central tendency risk. In general, one should avoid developing input distributions to a PRA model that yield a single risk distribution that intermingles, or represents both variability and uncertainty. By separately characterizing variability and parameter uncertainty, the output from a PRA will be easier to understand and communicate. A number of tools can aid in evaluating the uncertainty in estimated distributions for variability. Both simple and very complex approaches have been applied to this problem. Two basic methods for quantifying variability and parameter uncertainty simultaneously are described in Exhibit 1-4. PRAs that use these approaches can provide confidence bounds on percentiles of the risk distribution based on confidence bounds on one or more parameter estimates. Techniques for characterizing both variability and uncertainty in PRA are discussed in more detail in Chapters 4, 5, and 7, and Appendices D and E.

A common apprehension concerning the utility of PRA is that it may require too much knowledge to generate credible PDFs. In addition, if the inputs are not credible, then the outputs are not credible. These apprehensions are valid if the assessment attempts to model variability only. A risk assessor may feel that they can't specify a PDF because they don't know enough about the distribution or they are not sure that the values reported in the literature are representative of the site population of concern (i.e., *I am uncertain*). This is precisely the scenario when an analysis of uncertainty may be most helpful (see Chapter 3). While it is often true that a better quality and larger quantity of data would help to increase the confidence in risk estimates, uncertainty is not always a reason to avoid PRA. Indeed, when data are limited, an uncertainty analysis may help to inform the risk management decision process and may help in choosing the percentile of the risk distribution that corresponds with the RME (see Chapter 4).

Regarding uncertainty analysis, the use of probabilistic methods to propagate variability and uncertainty through risk models may have four key advantages over point estimate approaches:

1) Probabilistic methods can provide a more robust method of quantifying confidence in risk estimates than the point estimate approach. Monte Carlo simulation can be used to combine distributions of uncertainty for multiple input variables in a single simulation. By contrast, point estimate approaches combine point estimates of uncertainty in separate calculations, a technique that can yield estimates of plausible bounds for risk, but cannot yield an estimate of the upper and lower 95% confidence limits.

2) Techniques for sensitivity analysis that are available using PRA may help risk assessors to better identify the influential exposure factors and/or possible variations in the risk models used.
3) Probabilistic methods can account for correlations between input variables (e.g., body weight and skin surface area).

4) Probabilistic methods provide quantitative estimates of the expected value of information (EVOI) associated with obtaining more information (Morgan and Henrion, 1990). The importance of quantifying uncertainty in an EVOI framework is discussed in Appendix E.

Since both point estimate and probabilistic approaches in risk assessment are applied to the same conceptual models, uncertainties in the conceptual model are generally addressed in the same manner. If other models are available to explain or characterize a given phenomena, the risk estimates associated with each of those conceptual models could be compared to determine the sensitivity of the risk to the uncertainty in choice of model (see Chapter 2 and Appendix B). For example, when deciding on a contaminant concentration term for tetrachloroethylene in groundwater for a residential exposure assessment 10 years in the future, it would be appropriate to compare and contrast several fate and transport models and their results before deciding on a concentration term.

1.4 INTERPRETING THE RESULTS OF PRA

In PRA, risk managers are presented with distributions of estimated risks that explicitly consider variability and/or uncertainty in exposure. For ecological risk assessments, the risk distribution may also reflect variability and/or uncertainty in toxicity reference values (TRVs, see Chapter 5). In comparison with point estimate risk assessments, which generally provide numerical CTE and RME risk estimates to represent variability among individuals, PRA can provide the entire range of estimated risks as well as the likelihood of values within the range (i.e., the probability distribution). As noted above, the EPA Guidelines for Exposure Assessment (U.S. EPA, 1992a) states that the "high end" (or RME) of exposure for a population occurs between the 90th and 99.9th percentiles, with the 99.9th percentile considered a bounding estimate. Similarly, PRAs developed to support RME risk estimates for Superfund will also adopt this definition of a high-end risk.

- ***In general, risks corresponding to the 90th to 99.9th percentiles of the risk distribution may be considered plausible high-end risks. To support an evaluation of RME, a PRA generally should provide risk estimates covering this high-end range.***

This definition of the high-end range is based on a characterization of inter-individual variability in risk. In contrast to human health risk assessment, ecological risk assessments may focus on local population sustainability and community integrity. Because of the difference in the risk management objective, ecological risks are sometimes evaluated by assessing the risks to an average (rather than RME) individual in the population. However, as discussed in Section 5.1, inter-individual variability in ecological risk may be more important when there is a risk to threatened and endangered species. In general, uncertainty in risk estimates may be evaluated qualitatively or quantitatively, depending on the level of analysis that is warranted (i.e., using the tiered approach described in Section 1.6).

As discussed in Section 1.3.2, the Superfund Program generally seeks to establish risk-based cleanup goals that are protective for high-end exposed individuals (U.S. EPA, 1989a; 1990). The NCP (U.S. EPA, 1990) discusses a generally acceptable range for cumulative excess cancer risk of 10^{-6} to 10^{-4} for protecting human health. Furthermore, the NCP specifies 10^{-6} as a point of departure for determining

remediation goals when ARARs are not available or sufficiently protective. The range, 10^{-6} to 10^{-4}, will be referred to as the *risk range* in this guidance. If the Hazard Index (HI) is greater than 1, there may be a concern for potential non-cancer effects. Generally, where cumulative carcinogenic risk to an RME individual is less than 10^{-4}, and the non-carcinogenic hazard quotient is less than or equal to 1, action is not warranted unless there are adverse environmental impacts or ARARs (e.g., maximum contaminant levels) are exceeded (U.S. EPA, 1991b).

The results of a PRA will generally yield a probability distribution for risk (or Hazard Index), from which the risk corresponding to the CTE and RME can be identified. In general, a risk manager will identify an RME from the high end of the distribution (i.e., the 90th to 99.9th percentiles) for risk (or Hazard Index).

- *For clarity in this guidance, the 90th to 99.9th percentiles of the risk distribution are collectively referred to as the recommended RME range.*

Therefore, in order to utilize PRA results to establish that a cleanup goal is sufficiently protective, two questions will generally need to be addressed:

1) How will the RME risk be identified from the RME range of the risk distribution? and
2) How will information on uncertainty in the high-end risk estimates be utilized in this process?

Addressing these questions generally involves important risk management considerations, namely, the degree of protectiveness that should be factored into decisions about site cleanup goals. In utilizing PRA to establish remediation goals, risk managers may compare the RME range of risk estimates to the risk levels of concern. If, after considering uncertainty, and specifically, the potential for underestimating of risk, a risk level of concern falls outside the RME range, interpreting the results of a PRA is straightforward. For example, if the 90th percentile of the risk distribution is greater than the level of concern and uncertainty is low, there is strong evidence to support remedial action. Likewise, if the 99.9th percentile is below the level of concern and uncertainty is high, there is strong evidence to support a no-action decision. However, if the risk level of concern falls within the RME range, a closer look at site-specific factors may be warranted to determine whether the RME risk exceeds the level of concern. Site-specific information may be used to quantify uncertainty in the risk estimates as well as to identify the percentile of the risk distribution that appropriately represents the RME risk. As noted in Section 1.3.2, the RME risk identified from the results of a PRA is likely to differ from that of the point estimate approach, and either method can provide useful information for risk management.

In selecting a single percentile of the risk distribution as the basis for risk management decisions, risk managers will typically decide which percentile best represents the RME estimate at the site. The intent of this descriptor is to convey an estimate of risk in the upper end of the risk distribution, while providing a risk manager with the flexibility to select a different (higher or lower) percentile depending on site-specific information regarding exposure and toxicity (U.S. EPA, 1989a). In general, the percentile that appropriately represents the RME individual should be based on quantitative information and professional judgment. In particular, risk managers may need to understand what sources of variability and uncertainty are already explicitly accounted for by the modeling approach and inputs (i.e., point estimates and/or probability distributions) used to estimate the risk distribution. Greater detail on how to select an appropriate percentile from the RME range based on site-specific information is provided in Chapter 4.

- *In PRA, a recommended starting point for risk management decisions regarding the RME is the 95th percentile of the risk distribution.*

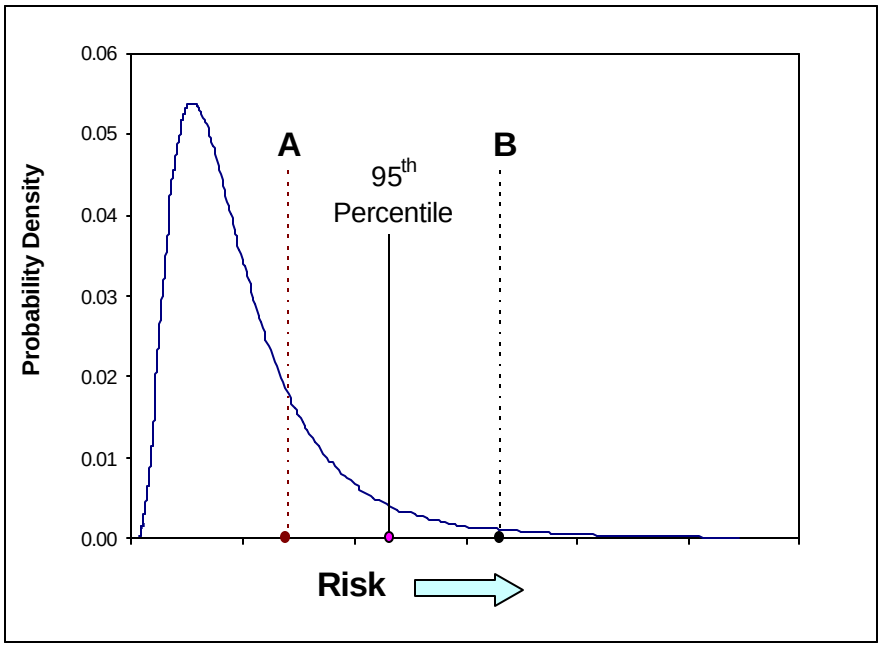


Figure 1-3. Example of a probability distribution for risk illustrating the 95th percentile and two different risk levels of concern (A and B). Assuming the 95th percentile corresponds to the RME, the need for remedial action depends on how the RME risk compares with the risk level of concern. For Case A (RME > level of concern), remedial action may be warranted. For Case B (RME < level of concern), remedial action may be unnecessary.

The 95th percentile of the risk distribution generally is an appropriate description of high-end exposure as identified by the Presidential/Congressional Commission on Risk Assessment and Risk Management in 1997. The 95th percentile can be used as a starting point for risk characterization when site-specific information is too limited to identify a more appropriate percentile. As shown in Figure 1-3, if the 95th percentile is greater than the level of concern, remedial action may be warranted. Conversely, if the 95th percentile is less than the level of concern, a no-action decision may be warranted. If a different percentile is identified from the RME range (90th to 99.9th percentile) to evaluate the RME risk, the same general approach would apply. The site-specific cleanup level is determined using the nine criteria specified in Section 300.430(e)(9)(iii) of the NCP.

1.5 ADVANTAGES AND DISADVANTAGES OF POINT ESTIMATE AND PROBABILISTIC APPROACHES

EXHIBIT 1-5
PRA MAY REQUIRE...

- additional time and funds to collect data;
- additional time and funds to conduct the assessment;
- specialized work plans;
- contractors with expertise in PRA to develop the assessment;
- additional time to discuss the assessment with management and stakeholders; and
- experts to review the results.

Both point estimate and probabilistic approaches can provide useful information for risk characterization. However, as discussed throughout this chapter, there are advantages and disadvantages associated with both methods that should be weighed carefully prior to choosing to conduct a PRA. A point estimate approach should generally be performed prior to considering a PRA. If there is a clear value added from performing a PRA, then the use of PRA as a risk assessment tool may be considered.

- *A point estimate approach is conducted for every risk assessment, and a probabilistic analysis may not always be needed.*

By relying on the full scope of available information, PRA can often provide a more complete characterization of risk, as well as a quantitative description of the uncertainties in risk estimates. However, PRA generally involves additional effort throughout the risk assessment process (see Exhibit 1-5) and may not needed for risk management at every site. Not all PRA's will involve the same level of effort to provide useful information for risk management decisions; the tiered approach given in Section 1.6 is recommended to determine the appropriate level of analysis. Also, the potential for misinterpretation of methods and conclusions is generally increased in PRA due to the greater complexity of the analysis. Some of the major advantages and disadvantages of both methods are summarized in Tables 1-1 and 1-2.

Compared to a point estimate risk assessment, a PRA typically involves more effort on the part of both the risk assessor and the risk manager (see Table 4-2). Given the constraints of PRA, it is important to determine what potential value may be added. An important question to ask is - *Does the benefit of PRA outweigh the cost and time involved in conducting a PRA*? For example, PRA may provide insight regarding both the decision to undertake active remediation and the extent to which remediation is necessary. For many sites, the additional information provided by a PRA will not affect the decision that would have been made with a point estimate approach alone. PRA will generally be most useful at complex sites where the decision whether to take action is unclear, and the stakes are high (both in terms of remediation costs and risks to human health and the environment).

EXHIBIT 1-6

STAKEHOLDERS POTENTIALLY INVOLVED IN DECISION-MAKING PROCESS FOR PRA

- EPA risk assessors and managers
- members of the public
- representatives from state or county environmental or health agencies
- other federal agencies (i.e., health agencies, NRDA trustees, etc.)
- tribal government representatives
- Potentially Responsible Parties (PRPs) and their representatives
- representatives from federal facilities (Department of Defense, Department of Energy, etc.)

1.6 A TIERED APPROACH FOR PRA

A tiered, or stepwise, approach to PRA is advocated, as shown in the flowchart illustrated in Figure 1-4. Tiered approaches to undertaking PRA have been discussed in the past (Brand and Small, 1995; Dakins et al., 1994; 1995; Finkel and Evans, 1987; Morgan and Henrion, 1990). In addition, tiered approaches are commonly used for ecological risk assessment (U.S. EPA, 1997a; 1998a). The level of analysis and sophistication of methods used to quantify variability and uncertainty in exposure and toxicity can vary in complexity depending on site-specific requirements. A tiered approach begins with a relatively simple analysis and progresses stepwise to more complex analyses. The level of complexity should match the site-specific risk assessment and risk management goals.

Decision points are those stages of the risk assessment process at which existing information is reviewed and decisions are made about next steps (Figure 1-3, diamonds). Some of these decision points may involve both scientific and policy considerations, and are analogous to *scientific/management decision points* (SMDPs) used in ecological risk assessment (U.S. EPA, 1997a), as presented in Chapter 5. SMDPs (indicated by ellipses in Figure 1-4) provide an opportunity for re-evaluation of direction and goals of the assessment at critical points in the process. Specific activities leading to these decision points are indicated by rectangles in Figure 1-4.

- *SMDPs are used to determine the next step of the process, and input from all stakeholders may be considered.*

Each decision point can be viewed as scoping for the next phase of the risk assessment. The risk manager may wish to hold a series of meetings to make decisions about the next step of a PRA involving the risk assessors and other appropriate personnel. Some potential stakeholders who may participate in discussions at decision points are listed in Exhibit 1-6.

PRA may involve highly technical concepts and discussions of PRA may tend to focus on the interchange between EPA and outside "experts". To counter balance such a tendency, managers generally should give particular emphasis to the inclusion of input from members of the general public who are (or may in the future be) exposed to contaminants from a site. A potential stakeholder may have an interest in the outcome of the remedial process. Communication between EPA and other stakeholders generally will continue throughout the risk assessment process. Often, stakeholders will possess valuable information that can improve the quality of the assessment (see Chapter 8).

1.6.1 BEGIN WITH SCOPING, PROBLEM FORMULATION, AND SCREENING ASSESSMENT

All risk assessments should begin with problem formulation and scoping (U.S. EPA, 1989a; 1997a). Problem formulation is generally an iterative process where substantial reevaluation may occur as new information and data become available, with the goal of achieving consensus on the problem formulation and analysis methods. An initial scoping meeting generally occurs prior to any risk assessment activities, as shown in the flowchart (Figure 1-4). The possibility that a PRA might be conducted is discussed at this early stage, taking into consideration the available information and the potential value added by quantifying variability or uncertainty. Also at this stage, a conceptual site model is developed that presents the contamination sources, contaminated media, plausible exposure pathways, and receptors at a specific site. During the initial scoping effort, additional factors that may be discussed by risk assessors and risk managers include: the extent of available data and potential needs for additional data collection; relevant exposure pathways (complete or incomplete); and pathways and variables that are expected to have a significant impact on the outcome of the risk estimates.

After scoping, the next step of the risk assessment process is to perform risk screening, considering all relevant site-specific pathways and using either preliminary remediation goals (PRGs) or risk-based concentrations (RBCs) calculated with default CTE and RME assumptions. After screening the concentrations for each contaminant for appropriate pathways, contaminants that exceed risk-based screening levels (contaminants of concern, COCs) will generally be evaluated further in a baseline risk assessment (Tier 1). For COC's that do not exceed regulatory criteria, the available data suggest that these COCs do not pose significant risks, and generally no remedial action need be taken. In general, however, consideration should be given to uncertainty (e.g., limited sampling data) or variability (e. g., particular patterns of site use) in factors that may suggest the potential for higher site risks. The level of analysis that is to be performed should generally be presented as discussed in RAGS Vol. 1, Part D (U.S. EPA, 1998b), and may include the following information: the extent of potential site remediation, degree of uncertainty associated with the exposure information available for each portion of the site conceptual model, and value added in the decision process. Documentation of the screening level analysis should be presented in a standard risk assessment format (RAGS Vol. 1 Part D) and should proceed through hazard identification and selection of chemicals of potential concern (U.S. EPA, 1989a; 1995b; 1996).

1.6.2 TIER 1 OF THE PRA

Figure 1-4, Steps A and B

The initial steps of every PRA will generally involve a point estimate risk assessment. If the point estimate(s) of risk are greater than the level of regulatory concern, a risk assessor (together with other stakeholders) may consider whether or not the existing information will support a remedial decision, or whether additional risk assessment activities are warranted. At several points in the tiered approach, a question is posed, "Are the Risks a Concern?". In order to address this question, a risk assessor (and

stakeholders) will generally consider the likelihood that the CTE and RME risk estimate exceeds a target risk level. It may also be important to consider the confidence in the risk estimates; that is, an RME risk estimate may be above or below a risk target, but judgment will be needed to determine the level of confidence that this risk estimate is sufficiently protective. The risk may be a concern if additional information on variability or uncertainty could lead to a different decision regarding remedial action. If additional probabilistic analysis is unlikely to make a difference in the risk management decision, then a decision generally should be made to not continue further with the tiered process for PRA.

1.6.3 TIER 2 OF THE PRA

Figure 1-4, Step C

Additional risk assessment activities should generally include an initial sensitivity analysis. This may be either a qualitative or quantitative analysis depending on the complexity of the risk assessment at this point. For example, incidental ingestion of soil by children is often an influential factor in determining risk from soil, a fact recognized by risk assessors. This recognition is a *de facto* informal sensitivity analysis. A quantitative sensitivity analysis can also be performed to identify those exposure variables with the greatest influence on risk estimate. Quantitative techniques for sensitivity analysis are presented in Chapter 2 and Appendix B.

Figure 1-4 , Step D

The decision to perform a PRA involves determining whether or not there are sufficient site-specific or representative data to describe variability and/or uncertainty in the exposure and toxicity variables used to estimate risk.

- *If there are data gaps for important exposure and toxicity variables, then PRA efforts should generally focus on quantifying the uncertainty in risk estimates.*

To make a decision to use PRA, there should be a clear benefit for at least one contaminant and one exposure pathway at the site. A sensitivity analysis (Step C) can be used to determine the factors that most strongly influence the risk estimate. Often, a small subset of exposure pathways and variables dominate the variability and uncertainty in risk estimates. Limiting the probabilistic descriptions to these influential factors is advisable, and will have only minor effects on the output distribution. If sufficient information is available to characterize variability and/or uncertainty in the key variables, an initial PRA may be performed (Step G).

Figure 1-4 , Steps E and F

A PRA will not, by itself, decrease the uncertainty of a risk estimate. The presence of significant data gaps may yield highly uncertaint risk estimates. After identifying important sources of variability and uncertainty, and comparing the point estimates of risk to level(s) of concern, a determination can be made regarding the adequacy of the available data for quantifying variability and uncertainty in risk. If significant data gaps exist, and reducing uncertainty in risk estimates is likely to assist in the risk management decision, additional data collection may be warranted. EPA guidance on data useability in risk assessment and the data quality objectives process should be considered whenever additional sampling or data evaluation is planned (U.S. EPA, 1992d; 1994b).

Table 1-1. Advantages and Disadvantages of Point Estimate Approach

ADVANTAGES	DISADVANTAGES
Uses RME input assumptions (i.e., combinations of central tendency and high-end values) to provide a plausible high-end estimate of risk.	Results in CTE and RME point estimates of risk, which may be viewed as "bright lines" that do not reflect uncertainty (i.e., risk levels are either above or below these risk estimates).
Useful as a screening method - may indicate that RME risks are either much greater or much less than a regulatory level concern.	Information from sensitivity analysis is generally limited to dominant exposure pathways and chemicals of concern; does not highlight the key exposure variables and uncertain parameters.
Central tendency and RME estimates of risk provide a semi-quantitative measure of variability.	Does not provide a measure of the probability that risk exceeds a regulatory level of concern, or the level of confidence in a risk estimate.
Employs a consistent approach and standardized reporting methods (U.S. EPA, 1998b).	Provides fewer incentives for collecting better or more complete information.
Easily understood and communicated.	May introduce inconsistency in risk estimates across sites due to different choices of point estimates.
Requires less time to complete; not as resource intensive.	May not utilize all available data for characterizing variability and uncertainty in risk estimates.

Table 1-2. Advantages and Disadvantages of Probabilistic Risk Assessment

ADVANTAGES	DISADVANTAGES
Can make more complete use of site data to characterize variability and uncertainty in risk.	Sufficient information may be lacking on variability and uncertainty for important exposure and/or toxicity variables.
Quantitative data on the uncertainty in exposure variables can be modeled and may support statistical confidence limits on risk estimates.	May require more time and resources to select and fit probability distributions.
Sensitivity analysis can identify the exposure variables, probability models, and model parameters that strongly influence estimates of variability and uncertainty in risk.	May convey false sense of accuracy unless the exposure models and probability distributions are representative of site conditions.
Puts the risk assessment in a *Value-of-Information* framework (see Appendix E). Can identify data gaps for further evaluation/data collection and can use wider variety of site-specific information.	May introduce inconsistency in risk estimates across sites due to different choices of probability distributions and risk percentile corresponding to the RME risk.
May provide more comprehensive summary of risk estimates, thereby supporting a more informed risk management decision process.	May require more time and resources to evaluate simulation results. Complex approaches may obscure important assumptions or errors in methodology.
Allows available site-specific information to inform the choice of high-end percentile from the risk distribution that corresponds with RME risk.	May require greater effort to communicate methodology and results.

Figure 1-4. Flow chart showing the progression and increasing complexity of a PRA.

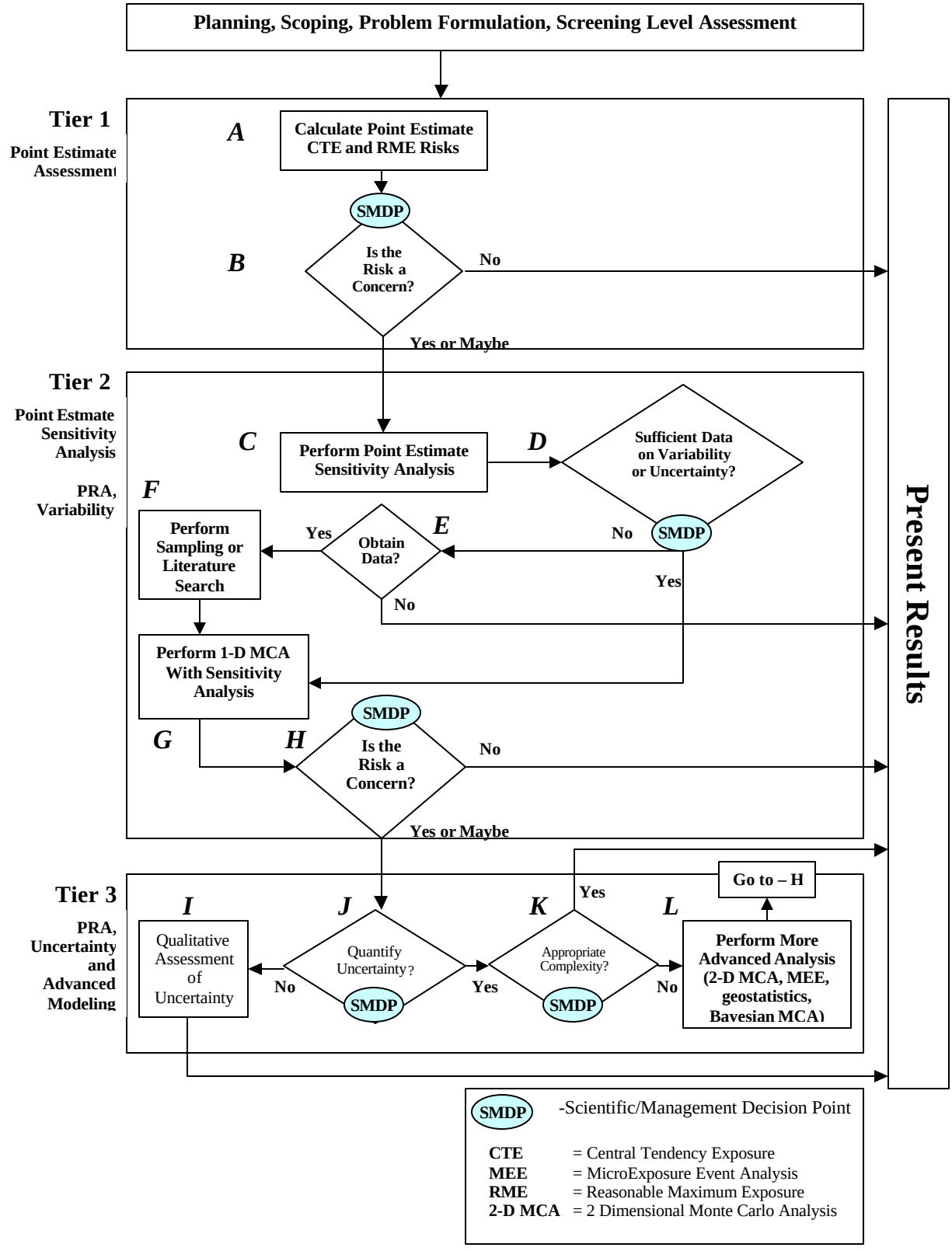


Figure 1-4 , Step G

If a PRA is advisable, then the first task is generally to perform an initial probabilistic simulation with a refined sensitivity analysis. The initial PRA is generally a 1-dimensional Monte Carlo analysis (1-D MCA) in which the probability distributions for selected exposure and toxicity variables represent inter-individual variability. Ideally, the probabilistic analysis will focus on the subset of exposure variables that have the greatest influence on the risk estimates. However, given the limitations of point estimate sensitivity analysis (see Appendix C), it may be difficult to identify a subset of exposure variables at the onset of the PRA. For multimedia exposure models, results of point estimate analyses may help to identify the contaminants and exposure pathways that dominate the aggregate risks. Sensitivity analyses may also be performed to identify the subset of parameters that dominate the uncertainty in risk estimates; this information can be used to focus the approaches for quantifying uncertainty in Tier 3.

- *A PRA will usually incorporate an iterative process in which the level and complexity of the analysis increases until the scope of the analysis matches the scope of the problem.*

The initial probabilistic simulation is the centerpiece of the workplan for the next tier of a PRA. Note that the risk assessment goals for ecological risk assessment may support a departure from Figure 1-4 at this point. A decision process specific to ecological risk assessment is presented in Chapter 5.

Figure 1-4 , Step H

The results of the initial PRA can be used to determine if the analyses in Tier 3 should be considered. In general, the basic question that should be addressed at this point is whether or not the risk estimates from both the point estimate approach and the 1-D MCA are of concern for the potentially exposed population. If the risk descriptor(s) (e.g., CTE and RME risks) are significantly greater than the level of concern, it may be prudent to stop here and report the results using standard reporting formats provided by RAGS Vol. 1, Part D (U.S. EPA, 1998b). As recommended in Section 3.1.3 of RAGS Vol. 1 Part D, the results of PRA (including the assessment of confidence and uncertainty) generally should be presented as part of the Risk Characterization portion of the Baseline Risk Assessment Report or as an appendix, in accordance with regional preferences. If the qualitative estimate of uncertainty is sufficiently great, a quantitative analysis of uncertainty provided in Tier 3 should be performed. As indicated by the SMDP in Figure 1-4, communication between risk assessors, risk managers, and stakeholders is important at this stage.

1.6.4 TIER 3 OF THE PRA

Figure 1-4, Steps I and J

If, after further consideration of available data and modeling approaches, a quantitative analysis of uncertainty is not considered useful or appropriate, then a qualitative discussion of uncertainties is generally sufficient for the PRA report.

Figure 1-4, Steps K and L

If uncertainty in important variables can be quantified, then modeling approaches that separately characterize variability and uncertainty should be considered. Examples include 2-D MCA, Microexposure Event Analysis, geostatistical analysis of concentration data, and Bayesian statistics.

These are discussed in detail in Appendix E. Several different approaches may be employed, in an iterative fashion, as results are obtained from each analysis.

1.7 GUIDING PRINCIPLES FOR CONDUCTING AN ACCEPTABLE PRA

In the EPA memorandum, *Use of Probabilistic Techniques in Risk Assessment* (U.S. EPA. 1997c), several conditions for conducting a scientifically sound PRA are outlined. A PRA should generally address these conditions to ensure that adequate supporting data and credible assumptions are used in the assessment. These conditions are as follows:

1. The *purpose and scope* of the assessment should be clearly articulated in a *"problem formulation"* section which should include a full discussion of any highly exposed or susceptible subpopulations evaluated (e.g., children, the elderly, etc.). The questions the assessment attempts to answer should be discussed and the assessment endpoints should be well defined. In accordance with RAGS Part A, the problem formulation should indicate that the risks are being evaluated in the absence of any remedial action. Both current and future risks at the site should be evaluated.

2. The *methods used* for the analysis (including all models used, all data upon which the assessment is based, and all assumptions that have a significant impact upon the results) should be documented and easily located in the report. This documentation should generally include a discussion of the degree to which the data used are representative of the population under study. Possible sources of bias inherent in the input distributions should be discussed along with the expected impacts on the resulting risk estimates. For example, if a site-specific study of fish consumption indicated consumption rates are five to ten times higher than other studies from similar populations, this possible bias or inaccuracy should be discussed in the document. Also, this documentation should include the names of the models and software used to generate the analysis. Computer programs should generally be described in sufficient detail to allow the reviewer to understand all aspects of the analysis. Computer code should provide adequate documentation and annotation. In summary, sufficient information should be provided to allow the results of the analysis to be independently reproduced.

3. The *results of sensitivity analyses* should be presented and discussed in the report. The more complex probabilistic techniques will typically be applied to the contaminants, pathways, and factors of importance to the assessment, as determined by sensitivity analysis or other basic requirements of the assessment.

4. The *presence or absence of moderate to strong correlations or dependencies* between the input variables should be discussed and accounted for in the analysis, along with the effects these have on the output distributions.

5. *Information for each input and output distribution* should be provided in the report. This includes tabular and graphical representations of the distributions (e.g., plots of probability density functions and/or cumulative distribution function plots) that indicate the location of any point estimates of interest (e.g., mean, median, 95^{th} percentile, 99.9^{th} percentile, etc.). The selection of distributions should be explained and justified. For both the input and output distributions, variability and uncertainty should be differentiated where possible.

6. The *numerical stability of the central tendency and the higher end* (i.e., upper tail) of the output distributions should be evaluated and discussed. For purposes of PRA, numerical stability refers to the observed numerical changes in parameters of the output distribution (e.g., median, 95th percentile) from a Monte Carlo simulation as the number of iterations increases. Because most risk equations are linear and multiplicative, distributions of risk will generally approximate a lognormal. In general, the tails of the distribution are less stable than the central tendency, and the rate of convergence for the tails will depend on the form of the risk model and the skewness of the probability distributions selected for input variables. Given the current speed of computers, numerical stability is generally not a concern for most 1-D MCA models; however, it can be an important consideration for more complex simulations, such as with 2-D MCA models.

7. *Calculations of exposures and risks using point estimate methods* should be reported. If results of PRA calculations differ substantially from point estimate calculations, a risk manager may benefit from understanding the reasons for the differences and the relative strengths of the different approaches. Sometimes, a closer look at uncertainties in the underlying data, assumptions, and models will lead a risk assessor to revisit certain assumptions in order to provide a more consistent basis for comparison. Furthermore, point estimates may be used to answer scenario-specific questions and to facilitate risk communication.

8. Since *exposure assumptions (e.g., exposure duration, body weight) are sometimes embedded in the toxicity metrics* (e.g., Reference Doses, Reference Concentrations, cancer unit risk values), the exposure estimates from the probabilistic output distributions may be aligned with the toxicity metric.

In addition, the following conditions specific to the Superfund program should generally be addressed:

1. For non-EPA lead PRAs, a work plan should be submitted for review and approval by the appropriate EPA regional office prior to submission of a PRA.

2. A tiered approach should be used to determine the level of complexity appropriate for the risk assessment. The decision to ascend to a higher tier of complexity should generally be made in conjunction with the risk manager, regional risk assessment personnel, and other stakeholders.

3. At this time, for human health risk assessments, toxicity values will generally be characterized by point estimates because of limitations in the data and techniques for characterizing toxicity to humans using distributions. Only if adequate supporting data are available to characterize variability or uncertainty in toxicity values will the Agency consider the use of PDFs to characterize toxicity. The Agency will determine the adequacy of supporting data on a case-by-case basis, pending consultation with EPA Headquarters (Office of Emergency and Remedial Response). For some sites, uncertainty in the toxicity values may be an important source of uncertainty in risk estimates. For ecological risk assessment, variability or uncertainty in toxicity values may be characterized as distributions.

REFERENCES FOR CHAPTER 1

Brand, K.P. and M.J. Small. 1995. Updating uncertainty in an integrated risk assessment: conceptual framework and methods. *Risk Anal.* 15(6): 719-730.

Dakins, M.E., J.E. Toll and M.J. Small. 1994. Risk-based environmental remediation: decision framework and role of uncertainty. *Environ. Tox. and Chem.* 13(12): 1907-1915.

Dakins, M.E., J.E. Toll and M.J. Small. 1995. Risk-based environmental remediation: Bayesian MC/P analysis and the expected value of sample information. *Risk Anal.* 16(1): 67-79.

Finkel, A.M. and J.S. Evans. 1987. Evaluating the benefits of uncertainty reduction in environmental health risk management. *J. of the Air Pollution Control Assoc.* 37: 1164-1171.

Finley, B.L. and D.J. Paustenbach. 1994. The benefits of probabilistic exposure assessment: Three case studies involving contaminated air, water and soil. *Risk Anal*. 14(1): 53-73.

Morgan, G.M. and M. Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge University Press. New York.

National Academy of Sciences (NAS). 1983. *Risk Assessment in the Federal Government: Managing the Process*. National Academy Press, Washington, DC.

National Academy of Sciences (NAS). 1994. *Science and Judgment in Risk Assessment.* National Academy Press, Washington, DC.

Presidential/Congressional Commission on Risk Assessment and Risk Management. 1997. *Risk Assessment and Risk Management in Regulatory Decision Making*. Final Draft, Volume 2.

U.S. EPA. 1989a. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment).* Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002. NTIS PB90-155581.

U.S. EPA. 1989b. *Risk Assessment Guidance for Superfund.* (RAGS)*: Volume II. Environmental Evaluation Manual*. Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/001.

U.S. EPA. 1990. National Oil and Hazardous Substances Pollution Contingency Plan. Final Rule. 40 CFR 300: *55 Federal Register*, 8666-8865, March 8.

U.S. EPA. 1991a. *Human Health Evaluation Manual, Supplemental Guidance: "Standard Default Exposure Factors"*. Office of Solid and Emergency Response, Washington, DC. OSWER Directive 9285.6-03.

U.S. EPA. 1991b. *Role of the Baseline Risk Assessment in Superfund Remedy Selection Decisions*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive 9355.0-30.

U.S. EPA. 1992a. *Final Guidelines for Exposure Assessment*. EPA/600/Z-92/001. *57 Federal Register*, 22888-22938, May 29.

U.S. EPA. 1992b. *Supplemental Guidance to RAGS: Calculating the Concentration Term*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive 9285.7-081.

U.S. EPA. 1992c. *Guidance on Risk Characterization for Risk Managers and Risk Assessors.* Memorandum from F. Henry Habicht II, Deputy Administrator. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive 9285.7-17.

U.S. EPA. 1992d. *Guidance on Data Useability in Risk Assessment. Part A.* Final. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive 9285.7-09A. NTIS PB92-96336.

U.S. EPA. 1994a. *Use of Monte Carlo Simulation in Risk Assessments.* Region 3, Hazardous Waste Management Division. Office of Superfund Programs, Philadelphia, PA. EPA/903/F-94/001.

U.S. EPA. 1994b. *Guidance for the Data Quality Objectives Process*, EPA QA/G-4. Final. September.

U.S. EPA. 1995a. *Use of Monte Carlo Simulation in Performing Risk Assessments* (Technical Section). Region 8, Hazardous Waste Management Division, Superfund Management Branch Technical Guidance, Denver, CO, RA-10.

U.S. EPA. 1995b. *Memorandum from Carol Browner on Risk Characterization*. Office of the Administrator, Washington, DC. February 22.

U.S. EPA. 1996. *Final Soil Screening Guidance, May 17, 1996. Soil Screening User's Guide*. Office of Solid Waste and Emergency Response, Washington, DC. EPA 540/R-96/018.

U.S. EPA. 1997a. *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments.* Interim Final. Environmental Response Team, Edison, NJ. EPA/540/R-97/006, OSWER Directive #9285.7-25, June.

U.S. EPA. 1997b. *Lognormal Distribution in Environmental Applications*. Office of Research and Development, and Office of Solid Waste and Emergency Response, Washington, DC. EPA/600/R-97/006. December.

U.S. EPA. 1997c. *Memorandum from Deputy Administrator Fred Hansen on the Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment, and Guiding Principles for Monte Carlo Analysis.* Office of Research and Development, Washington, DC. EPA/630/R-97/001. May.

U.S. EPA. 1997d. *Exposure Factors Handbook, Volume 1.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fa.

U.S. EPA. 1997e. *Exposure Factors Handbook, Volume 2.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fb.

U.S. EPA. 1997f. *Exposure Factors Handbook, Volume 3.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fc.

U.S. EPA. 1998a. *Guidelines for Ecological Risk Assessment*. Final. National Center for Environmental Assessment, Washington, DC. EPA/630/R-95/002F.

U.S. EPA. 1998b. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part D, Standardized Planning, Reporting, and Review of Superfund Risk Assessments).* Interim Final. EPA/540/R-97/033. OSWER Directive # 9285.7-01D, PB97-963305.

U.S. EPA. 1999. *Ecological Risk Assessment and Risk Management Principles for Superfund Sites*. Final. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive # 9285.7-28P.

CHAPTER 2

SENSITIVITY ANALYSIS: HOW DO WE KNOW WHAT'S IMPORTANT?

2.0 INTRODUCTION

Sensitivity analysis, as it is applied to risk assessment, is an approach that utilizes a variety of mathematical and statistical techniques to determine which factors in a risk model influence risk most strongly. It provides a means of exploring, in a quantitative manner, the effect of a variety of "what-if" scenarios on the risk estimates. Sensitivity analysis can provide insight into the importance of selecting a particular model, including or excluding specific exposure pathways, and making certain assumptions with respect to model input parameters (U.S. EPA, 1997). Chapter 2 focuses on a set of graphical and statistical techniques that can be used to evaluate which variables in the risk model contribute most to the variation in estimates of risk. This variation in risk could represent variability, uncertainty, or both, depending on the type of risk model and characterization of input variables.

Sensitivity analysis is used in both point estimate and probabilistic approaches. The basic approach is to allow for a subset of the input variables to vary within prescribed ranges and to determine how much the model output changes in response to changes in the values for each input variable.

As shown in the tiered approach in the flowchart in Figure 1-4, a sensitivity analysis is useful at multiple steps of a probabilistic risk assessment (PRA). It may be performed several times for a single risk assessment to both guide the complexity of the analysis and communicate important results. The initial sensitivity analysis (Tier 2, Step C) is conducted with a point estimate risk assessment. It highlights which exposure pathways and variables most strongly influence the risk estimate and may provide useful information by being advanced through the PRA tiered approach. It can also help to determine if additional data collection and/or research efforts are warranted (Tier 2, Steps E and F). The next sensitivity analysis (Tier 2,Step G) is performed with a 1-D MCA to determine which variables have the greatest contribution to the variance in risk estimates. In ecological PRA (see Chapter 5), for which parameter uncertainty may be characterized using 1-D MCA, sensitivity analysis can be used to determine which parameters have the greatest contribution to the uncertainty in the risk estimates. Sensitivity analysis can also highlight important assumptions regarding the choice of probability distributions and truncation limits. Results from 1-D MCA may support a decision to explore more advanced modeling approaches in Tier 3.

- *A sensitivity analysis generally should be performed following both a 1-dimensional and 2-dimensional Monte Carlo analysis (1-D MCA and 2-D MCA) to determine which exposure assumptions and inputs contribute the most to variability and uncertainty, respectively.*

As noted in Chapter 1, MCA is one technique used in PRA, and is the focus of this guidance. An input variable contributes significantly to the output risk distribution if it is both highly variable *and* the variability propagates through the algebraic risk equation to the model output (i.e., risk). Changes to the

1

EXHIBIT 2-1

DEFINITIONS FOR CHAPTER 2

Continuous Variables - A random variable that can assume any value within an interval of real numbers (e.g., body weight).

Correlation - A quantitative expression of the statistical association between two variables; usually represented by the Pearson correlation coefficient for linear models, and the Spearman rank correlation coefficient (see below) for nonlinear models.

Discrete Variables - A random variable that can assume any value within a finite set of values (e.g., number of visits to a site in one year) or at most a countably infinite set of values.

Local Sensitivity Analysis - Evaluation of the model sensitivity at some nominal points within the range of values of input variable(s).

Monte Carlo Analysis (MCA) or Monte Carlo Simulation - The process of repeatedly sampling from probability distributions to derive a distribution of outcomes. MCA is one of several techniques that may be used in PRA.

Multiple Regression Analysis - A statistical method that describes the extent, direction, and strength of the relationship between several (usually continuous) independent variables (e.g., exposure duration, ingestion rate) and a single continuous dependent variable (e.g., risk).

Nonparametric Tests - Statistical tests that do not require assumptions about the form of the population probability distribution.

Range Sensitivity Analysis - Evaluation of the model sensitivity across the entire range of values of the input variable(s).

Sensitivity Analysis - Sensitivity generally refers to the variation in output of a model with respect to changes in the values of the model's input(s). Sensitivity analysis attempts to provide a ranking of the model inputs based on their relative contributions to model output variability and uncertainty. Common metrics of sensitivity include:

< Pearson Correlation Coefficient - A statistic r that measures the strength and direction of linear association between the values of two quantitative variables. The square of the coefficient (r^2) is the fraction of the variance of one variable that is explained by least-squares regression on the other variable.

< Sensitivity Ratio - Ratio of the change in model output per unit change in an input variable; also called *elasticity*.

< Spearman Rank Order Correlation Coefficient - A "distribution free" or nonparametric statistic r that measures the strength and direction of association between the ranks of the values (not the values themselves) of two quantitative variables. See Pearson (above) for r^2.

distribution of a variable with a high sensitivity could have profound impact on the risk estimate, whereas even large changes to the distribution of a low sensitivity variable may have minimal impact on the final result. This is important when trying to determine where to focus additional resources.

The benefits of sensitivity analysis applied to PRA are presented in Section 2.1. Details regarding common statistical techniques for conducting a sensitivity analysis and interpreting the results are given in Section 2.2 and in Appendix B.

2.1 BENEFITS OF SENSITIVITY ANALYSIS

Sensitivity analysis is beneficial to both risk assessors and risk decision makers for a number of reasons. Exhibit 2-2 highlights the types of information that sensitivity analysis can add to a Superfund risk assessment. As discussed in Section 2.0, a sensitivity analysis applied in a tiered approach can be useful in deciding which exposure pathways and assumptions are carried forward from a point estimate risk assessment into a 1-D or 2-D MCA. By identifying the variables that are most important in determining risk, one can also decide whether point estimates, rather than PDFs, can be used with little consequence to the model output (thereby reducing the level of effort associated with developing PDFs for all input variables). This information is important not only for designing 1-D MCA models of variability, but also for designing more complex analyses of uncertainty discussed in Appendix E (e.g., 2-D MCA models, geostatistical analysis, Bayesian analysis).

EXHIBIT 2-2

BENEFITS OF SENSITIVITY ANALYSIS

- **Uncertainty analysis** - *e.g., After quantifying parameter uncertainty, we are 95 percent confident that the RME risk is below the risk level of concern.*
- **PRA model design** - *e.g., How does my selection of a beta distribution over a lognormal distribution influence the 95th percentile of the risk distribution?*
- **Resource allocation** - *e.g., Two of the 10 exposure variables contribute 90 percent of the uncertainty in the RME risk estimate.*
- **Risk communication** - *e.g., For input variable X, if we were to use a distribution based on site-specific data instead of a national survey, we would expect a minimal change in the RME risk estimate.*

A hypothetical example showing results of a sensitivity analysis for a 1-D MCA is presented in Figure 2-1. For this example, hazard index (HI) is calculated using Equation 2-1 and the inputs given in Table 2-1.

$$HI = \frac{C \cdot I \cdot ED \cdot EF}{BW \cdot AT} \cdot \frac{1}{RfD} \qquad \text{Equation 2-1}$$

Five exposure variables are used to characterize variability in HI associated with an occupational soil ingestion pathway. HI is predominantly sensitive to water ingestion rate (r^2 = 88%) and, to a lesser

extent, body weight (r^2 = 7%) and exposure frequency (r^2 = 5%). The metric of sensitivity in this case is the square of the Spearman Rank Correlation Coefficient (see Appendix B.2.5). If the model output variable (e.g., HI) and input variable are highly correlated, it means that the input variable has an impact on the model. As shown in Figure 2-1 (bottom panel), to determine if the correlation is positive or negative, the correlation coefficient should not be squared. For risk equations in general, variables in the numerator of the equation (concentration, ingestion rate, exposure duration, etc.) will tend to be positively correlated with risk, while variables in the denominator (body weight, RfD for ecological receptors) will tend to be negatively correlated with risk. The greater the absolute value of the correlation coefficient, the stronger the relationship.

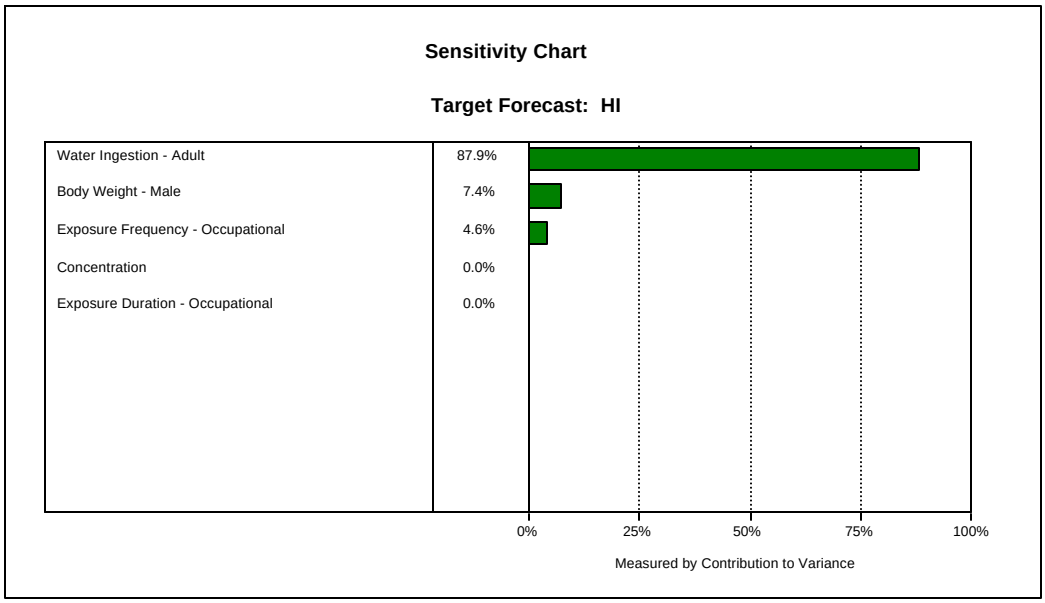


Figure 2-1 (1 of 2). Top panel - bar graph showing the r^2 values (square of Spearman rank correlation coefficient), a metric for the dependence of Hazard Index (HI) on exposure factors based on 1-D MCA for variability.

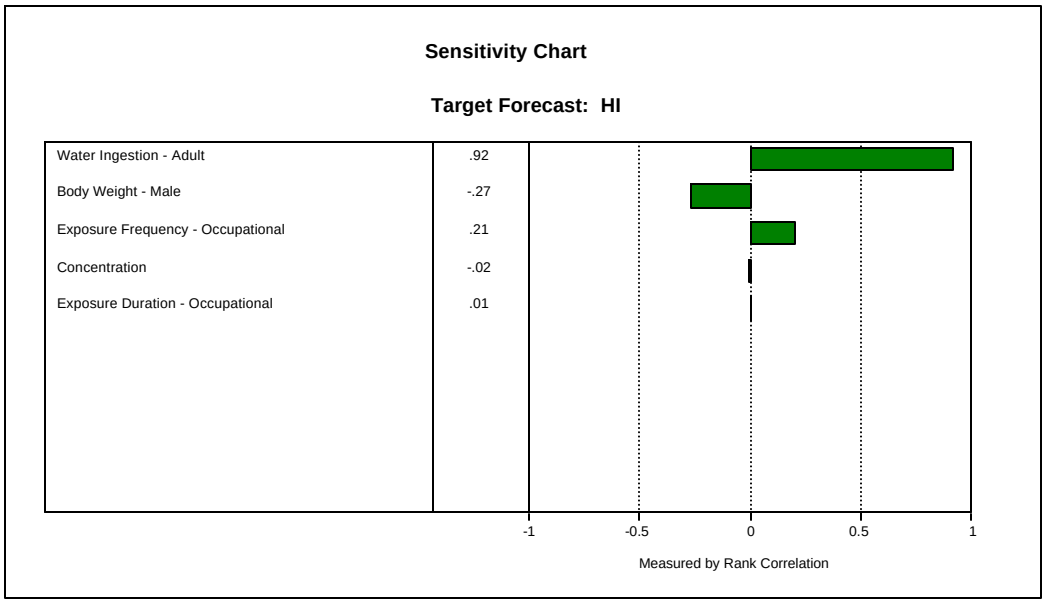


Figure 2-1 (2 of 2). Bottom panel - bar graph, sometimes referred to as "tornado plot", showing rank correlation coefficient. This graph is effective for showing both the relative magnitude and direction of influence (positive or negative) for each variable. In this example, the variable with the greatest effect on HI is the water ingestion rate. Concentration does not influence variability because, in this example, long-term average concentration is characterized by a point estimate (i.e., 95% UCL), rather than a probability distribution. Exposure duration does not influence variability because variability in ED is expressed in both the numerator (ED) and denominator (AT = ED x 365 for non-carcinogenic effects), and cancels out. Output was generated with Crystal Ball, which calculates the contribution to variance by squaring the rank correlation coefficient and normalizing to 100%.

Table 2-1. Input variables used in hypothetical example of hazard index associated with occupational exposure via water ingestion.

Variable	Distribution	Parameters	Units
Concentration in Water (C)	95% UCL	40	mg/L
Tap Water Ingestion Rate (I)	lognormal[1]	[1.3, 0.75]	L/day
Exposure Duration (ED)	empirical[2]	see below	years
Exposure Frequency (EF)	triangular	[180, 250, 350]	days/yr
Body Weight (BW)	lognormal[1]	[74.6, 12.2]	kg
Averaging Time (AT)	empirical[3]	see below	days
RfD	point estimate	0.5	mg/kg-day

[1] Parameter of lognormal distribution are [arithmetic mean, standard deviation].

[2] Parameters of empirical distribution for ED ~ [min, max, {x}, {p}] = [0, 30, {0.08, 0.18, 0.30, 0.44, 0.61, 0.84, 1.17, 1.72, 3.1, 6.77, 14.15, 23.94}, {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.975, 0.99}]

[3] AT = ED x 365 for non-carcinogenic risks (hazard index).

In this example, concentration has a correlation of 0 because a point estimate (i.e., 95% UCL) was used in the risk equation. Point estimates do not vary in a Monte Carlo simulation and, therefore, do not contribute to the variance in the output. This result does not mean that concentration is an unimportant variable in the risk assessment. Concentration may still contribute greatly to the uncertainty in the risk estimate. A sensitivity analysis of parameter uncertainty in a risk equation can be explored using iterative simulations, such as with 2-D MCA.

Exposure duration also contributes 0% to variance in HI, despite the fact that ED is characterized by an empirical distribution function. For this example, we are focusing on the non-cancer health effect associated with a contaminant in groundwater, and the averaging time (AT) is a function of the exposure duration (ED x 365). Therefore, algebraically, ED cancels out, and does not contribute to the variance in HI. In contrast, for evaluations of carcinogenic risks, the averaging time is generally 70 years (i.e., a constant), so ED would contribute to the variance in risk.

Decisions regarding allocation of future resources and data collection efforts to reduce lack of knowledge generally should take into consideration the most influential input factors in the model, and the cost of gaining new information about the factors. Sensitivity analysis is a key feature of determining the expected value of information (EVOI) (see Chapter 3 and Appendix E). Once a sensitivity analysis is used to identify an input variable as being important, the source of it's variability generally should be determined. If an input variable has a significant uncertainty component, further research and/or data collection can be conducted to reduce this uncertainty. Reducing major sources of uncertainty, such as the most relevant probability model for variability or the parameter estimates for the model, will generally improve confidence in the model output, such as the estimated 95th percentile of the risk distribution. A variable may contribute little to the variability in risk, but greatly to the uncertainty in risk (e.g., the

concentration term). Likewise, a variable may contribute greatly to the variance in risk, but, because the data are from a well characterized population, the uncertainty is relatively low (e.g., adult tap water ingestion rate).

An example of the output from a 2-D MCA of uncertainty and variability (see Appendix E) is shown in Figure 2-2. This 2-D MCA shows wide confidence limits around the 95th percentile of the distribution for risk. Assume for this example that the decision makers choose the 95th percentile risk as the RME risk. These wide confidence limits suggest that there is high uncertainty associated with this particular risk estimate. Next, assume a sensitivity analysis is run to identify the source(s) of that uncertainty, and the results, as shown in Figure 2-2, indicate that the soil concentration variable contributes most to the uncertainty in the 95th percentile risk estimate. Since both the sample size and variance impact the magnitude of the confidence limits for an arithmetic mean soil concentration, one way to reduce the confidence limits (and thus the uncertainty) would be to collect additional soil samples. Increasing the sample size (from n = 25 to n = 50) reduced the 90% confidence limits for the 95th percentile risk to below $1x10^{-5}$.

Although the uncertainty in a risk estimate can be reduced by further data collection if the sensitive input distribution represents uncertainty, this is not necessarily true for input distributions that represent variability. For example, variability in the distribution of body weights can be better characterized with additional data, but the coefficient of variation (i.e., standard deviation divided by the mean) may not be reduced.

Even if additional data collection efforts prove to be infeasible, identifying the exposure factors that contribute most to risk or hazard may be useful for risk communication. For example, assume that the input for exposure frequency has the strongest affect on the risk estimate for a future recreational open space. Further examination of this exposure variable reveals that the wide spread (i.e., variance) of the PDF is a result of multiple users (e.g., mountain bikers, hikers, individuals who bring picnics, etc.) of the open space who may spend very different amounts of time recreating. As a result of this analysis, the decision makers and community may decide to focus remediation efforts on protecting the high-risk subpopulation that is expected to spend the most time in the open space.

After determining which contaminants, media, and exposure pathways to carry into a PRA, numerical experiments generally should be performed to determine the sensitivity of the output to various distributions and parameter estimates that may be supported by the available information. Variables that do not strongly affect the risk estimates generally should be characterized with point estimates. This guidance document does not recommend a quantitative metric or rule of thumb for determining when a variable strongly affects the output; this must generally be determined on a case-by-case basis. A qualitative or quantitative analysis may be used depending on the complexity of the risk assessment at this point. For example, incidental ingestion of soil by children is often an influential factor in determining risk from soil, a factor recognized by risk assessors. This recognition is a *de facto* informal sensitivity analysis. An array of quantitative techniques are also available, ranging from something as simple as comparing the range of possible values (i.e., maximum - minimum) for each variable, to more complex statistical methods such as multiple regression analysis. Several of these methods are discussed in more detail in this chapter and in Appendix B.

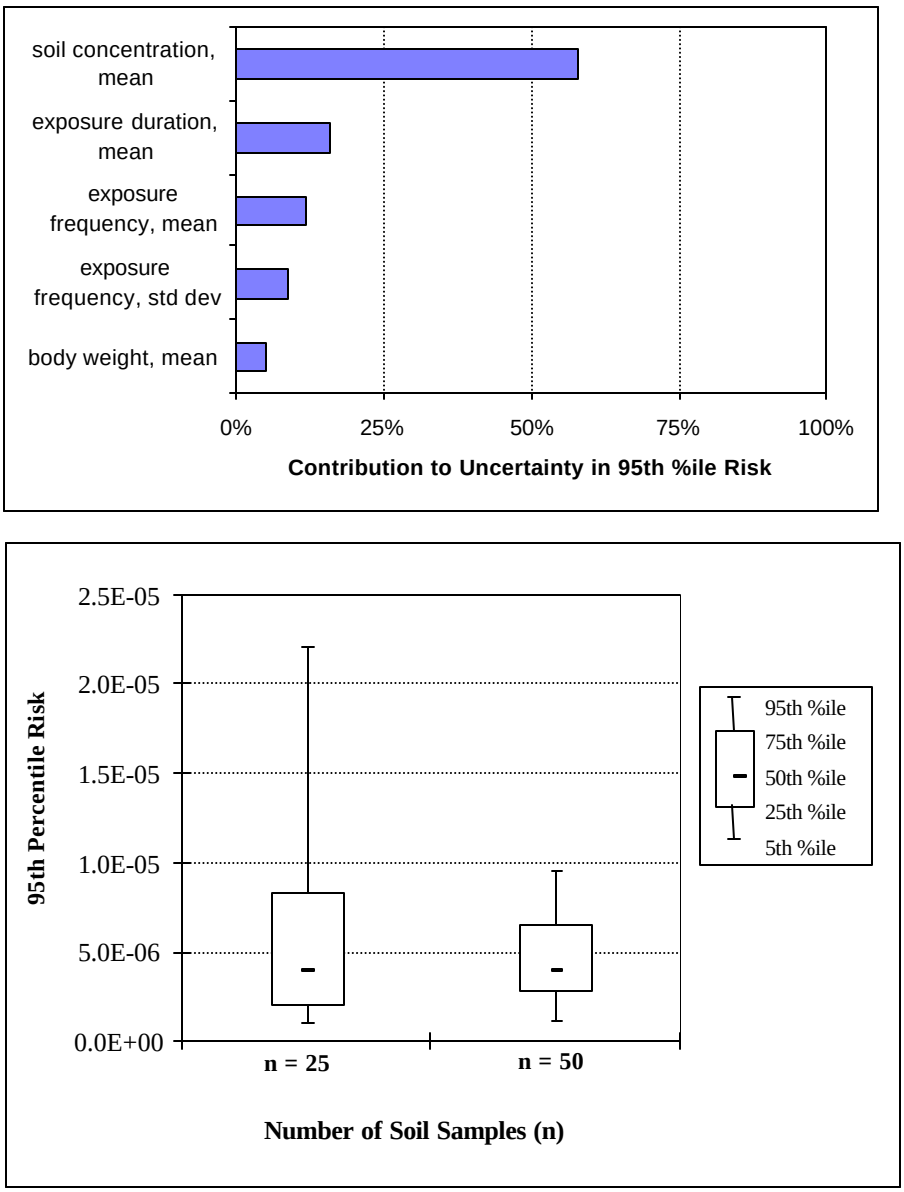


Figure 2-2. Results of 2-D MCA in which parameters of input distributions describing variability are assumed to be random values. Results of a sensitivity analysis (top graph) suggest that more than 50% of the uncertainty in the 95th percentile of the risk distribution is due to uncertainty in the arithmetic mean concentration in soil. The bottom graph gives box-and-whisker plots for the 95th percentile of the risk distribution associated with Monte Carlo simulations using different sample sizes (n = 25 and n = 50). For this example, the whiskers represent the 5th and 95th percentiles of the distribution for uncertainty, otherwise described as the 90% confidence interval (CI). For n = 25, the 90% CI is [1.0E-06, 2.2E-05]; for n = 50, the 90% CI is reduced to [1.2E-06, 9.5E-06]. While increasing n did not change the 50th percentile of the uncertainty distribution, it did provide greater confidence that the 95th percentile risk is below 1x10^{-5}.

2.2 COMMON METHODS OF SENSITIVITY ANALYSIS

Of the numerous approaches to sensitivity analysis that are available (see Exhibit 2-3), no single approach will serve as the best analysis for all modeling efforts. The best choice for a particular situation will depend on a number of factors, including the nature and complexity of the model and the resources available. A brief description of two of the more common approaches is provided in this chapter. Appendix B presents a more detailed discussion about these and other methods. Sensitivity analysis need not be limited to the methods discussed in this guidance, which focus on the more common approaches. A large body of scientific literature on various other methods is available (e.g., Hamby, 1994; Iman et al., 1988, 1991; Morgan and Henrion, 1990, Rose et al., 1991; Saltelli and Marivort, 1990; Shevenell and Hoffman 1993; U.S. EPA, 1997). Any method used, however, generally should be documented clearly and concisely. This documentation generally should include all information needed by a third party to repeat the procedure and corroborate the results, such as the exposure pathways and equations; a table with the input variables with point estimates, probability distributions and parameters; and tables or graphs giving the results of the sensitivity analysis and description of the method used.

EXHIBIT 2-3
INDICES OF SENSITIVITY ANALYSIS

- Local and/or range sensitivity ratios (i.e., elasticity)
- Normalized partial derivative
- Simple correlation coefficient (or coefficient of determination, r^2)
- Partial correlation coefficient / rank correlation coefficient
- Rank correlation coefficient
- Normalized multiple regression coefficient

2.2.1 GRAPHICAL TECHNIQUES

Simple scatter plots of the simulated input and output (e.g., risk vs. exposure frequency, or risk vs. arithmetic mean soil concentration) can be used to qualitatively evaluate influential variables. A "tight" scatter plot (i.e., a high r^2 value) suggests that a variable may significantly influence the variance in risk. Hypothetical scatter plots that may be used to identify sensitive and insensitive variables are shown in Figure 2-3. Another method for visualizing the relationship between all of the inputs and outputs is to generate a scatterplot matrix (Helsel and Hirsch, 1992). This graphic shows both histograms and scatterplots for all variables on the same page.

2.2.2 SENSITIVITY RATIOS: A SCREENING TOOL

Sensitivity ratios (SR) can be used for sensitivity analysis in both point estimate and probabilistic risk assessment. The approach is easy to understand and apply. The ratio is equal to the percentage change in output (e.g., risk) divided by the percentage change in input for a specific input variable (see Appendix B, Equation B-1). Risk estimates are considered most sensitive to input variables that yield the highest ratios. For simple exposure models in which the relationship between exposure and risk is linear, the ratio offers little information regarding the relative contributions of each input variable to the variance in risk, as demonstrated in Exhibit 2-4. However, for more complex models with nonlinear relationships

between inputs and outputs (e.g., environmental fate and transport models, pharmacokinetic models), the ratio can offer a useful screening tool to identify potentially influential input variables.

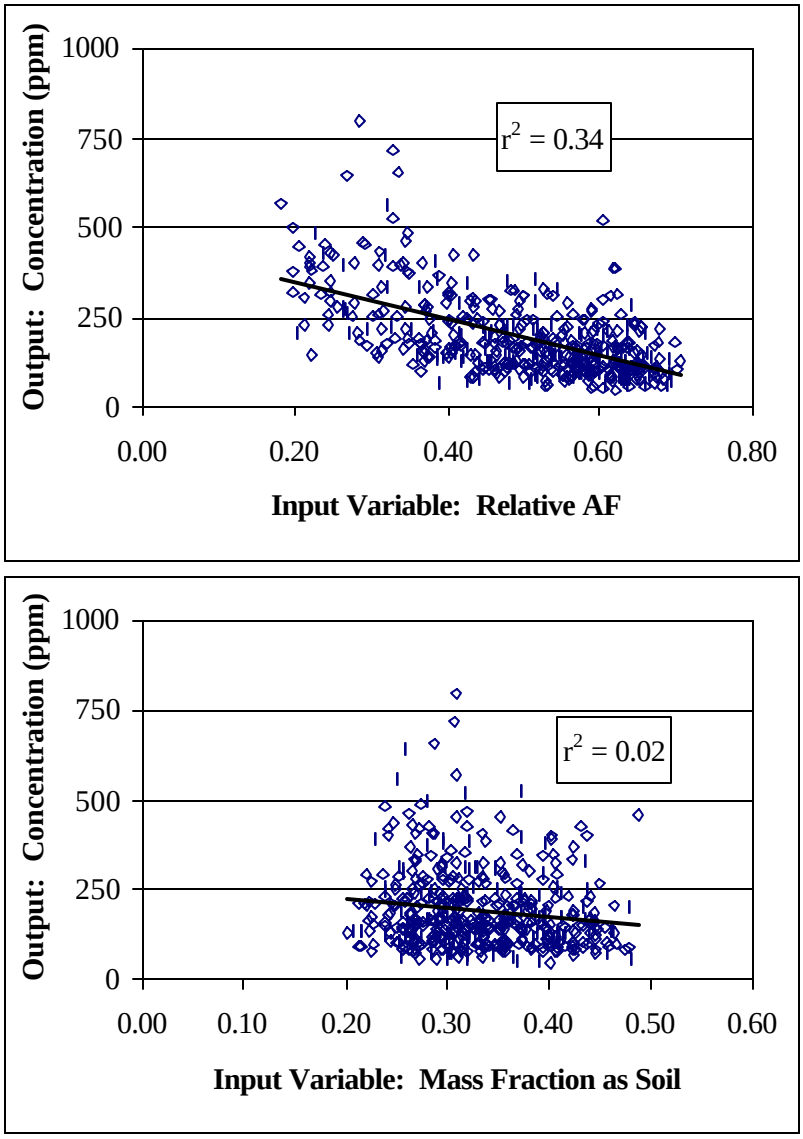


Figure 2-3. Scatterplots of simulated random values from a 1-D MCA of variability. The output from the model is a contaminant concentration in soil (C) that corresponds with a prescribed (fixed) level of risk for a hypothetical population (based on Stern, 1994). For each iteration of a 1-D MCA simulation, random values were simultaneously selected for all model variables and the corresponding concentration (C) was calculated. Inputs were simulated as independent random variables. Scatterplots of 500 consecutive random values and estimates of C are shown for two input variables: relative absorption fraction, RAF (top graph); and mass fraction of dust as soil, F (bottom graph). There is a moderate, indirect relationship between C and RAF (r^2 = 0.34), compared with the weak relationship between C and F (r^2 = 0.02), suggesting that the model output (C) is more sensitive to variability in RAF than F.

EXHIBIT 2-4

EXAMPLE OF SENSITIVITY RATIO CALCULATION

Using the risk model described by Equation 2-1, and hypothetical point estimates given below, sensitivity ratios were calculated for each of the exposure variables by modifying the point estimates ± 50% (i.e., 1/) = ± 2.0). Applying Equation B-3 from Appendix B yields the following results given by Table 2-2 below.

Table 2-2. Calculation results for sensitivity ratios of a linear equation for hazard index (HI).

Exposure Variable	**Point Estimate**, [-50%, +50%]	$\frac{(HI_2 - HI_1)}{HI_1}$	**Sensitivity Ratio (SR)**
Conc. in Water (C)	40, [20, 60]	[-50%, 50%]	[1.0, 1.0]
Tap Water Ingestion Rate (I)	1.30, [0.65, 1.95]	[-50%, 50%]	[1.0, 1.0]
Exposure Duration (ED)	0.61, [0.31, 0.92]	[0%, 0%]	[0.0, 0.0]
Exposure Frequency (EF)	250, [125, 365]	[-50%, 46%]	[1.0, 1.0]
Body Weight (BW)	74.6, [37.3, 111.9]	[100%, -33%]	[-2.0, -0.67]

[1]For non-cancer risks, AT = ED x 365; therefore, ED cancels out of the HI equation algebraically and does not contribute to the sensitivity ratio.
[2]Maximum EF = 365 days; therefore, maximum % change of point estimate = 46%.

The SR results suggest that relative affects of the input variables on HI cannot be readily distinguished. HI is equally sensitive to changes in ***C***, ***I***, and ***EF*** (SR = 1.0). In addition, for these variables, positive and negative changes to the point estimate yielded the same SR. HI is inversely proportion to changes in ***BW*** (i.e., as BW decreases, HI increases) and would appear to be most sensitive to this variable as ***BW*** decreases below the hypothetical default value of 74.6 kg. In contrast, the probabilistic sensitivity analysis results (Figure 2-1) suggests that ***I*** is by far the most influential variable. While the point estimate sensitivity analysis may be useful as a screening tool (e.g., identifying important variables in nonlinear equations, or important exposure pathways that contribute to aggregate risk estimates), in general, SR is not a robust method for identifying the major sources of variability and uncertainty in a risk model.

Sensitivity ratios can generally be grouped into two categories: local and range. For the local sensitivity ratio method, an input variable is changed by a small amount, usually ±5% of the nominal (default) point estimate, and the corresponding change in the model output is observed. For the range sensitivity ratio method, an input variable is varied across the entire range (plausible minimum and

maximum values). If local and range sensitivity results are different, the risk assessor can conclude that different exposure variables are dominating risk near the high-end (i.e., extreme tails of the risk distribution) than are dominating risk at the central tendency. This situation is likely to occur when there are nonlinear relationships between an input and output variable. Equation B-3 in Appendix B can be used to evaluate SR for different types of exposure models in which the intake equation is generally expressed as a simple algebraic combination of input variables.

2.2.3 SENSITIVITY ANALYSIS AND THE MONTE CARLO METHOD

Probabilistic sensitivity analysis can be generally grouped into two categories: 1) methods applied after a Monte Carlo simulation is run using the entire set of probability distributions for variables and unknown parameters; and 2) methods applied by running simulations in which different subsets of variables and/or uncertain parameters are assigned distributions, while all other inputs are set to their central values. This guidance focuses on the first category; further information regarding both techniques is given by Cullen and Frey (1999).

EPA anticipates that Monte Carlo analysis will be used for the majority of PRAs received by the Agency. In Monte Carlo analysis, the probability distributions assumed for the various input variables are used to generate a sample of a large number of points. Statistical methods are applied to this sample to evaluate the influence of the inputs on the model output. A number of different "indices" of sensitivity can be derived from the simulated sample to quantify the influence of the inputs and identify the key contributors. Most of these are based on an assumption that the model output *Y* varies in a monotonic, linear fashion with respect to various input variables (X_1, X_2, etc.). For example, an estimate of average daily intake (mg/kg-day) from multiple exposure pathways is linear with respect to the intake from each pathway. Since most risk models are linear with respect to the input variables, the output distribution (particularly its upper percentiles) tends to be dictated by the input variables with the largest coefficient of variation (CV), or the ratio of the standard deviation to the mean. For example, Equation 2-2 represents a simple expression for intake rate as a function of random variables X_1 and X_2 :

$$Y = X_1 + X_2 \qquad \text{Equation 2-2}$$

where X_1 and X_2 may represent dietary intake associated with prey species 1 and 2, respectively. If the same probability distribution was used to characterize X_1 and X_2, such as a normal distribution with an arithmetic mean of 100 and standard deviation of 50 (i.e., CV = 50/100 = 0.5), each variable would contribute equally to variance in Y. If, however, X_2 was characterized by a normal distribution with an arithmetic mean of 100 and standard deviation of 200 (i.e., CV = 200/100 = 2.0), we would expect Y to be more sensitive to X_2. That is, X_2 would be a greater contributor to variance in Y.

While the coefficient of variation may be a useful screening tool to develop a sense of the relative contributions of the different input variables, a common exception is the case when X_1 and X_2 have different scales. For example, Equation 2-3 is an extension of Equation 2-2:

$$Y = a_1 X_1 + a_2 X_2 \qquad \text{Equation 2-3}$$

where a_1 and a_2 are constants that may represent the algebraic combination of point estimates for other exposure variables. If $a_1 >> a_2$, then X_1 would tend to be the dominant contributor to variance, regardless of the CV for X_2.

The most influential random variables generally have the highest degrees of skewness or are related to the output according to a power function (Cullen and Frey, 1999). For example, Equation 2-4 presents an extension of Equation 2-3 in which there is a power relationship between X_2 and Y. In this example, assume Y represents the total dietary intake rate of cadmium for muskrats, X_1 and X_2 represent the dietary intake rate associated with prey species 1 and 2, respectively, a_1 and a_2 represent additional point estimates in the equation, and 2 is the power exponent. In general, for $2 > 1$, the total dietary intake rate (Y) will be more sensitive to the intake rate associated with species 2 (X_2) than species 1.

$$Y = a_1 X_1 + a_2 X_2^q \qquad \text{Equation 2-4}$$

Various statistical techniques, known collectively as regression analysis, can be applied to a linear equation to estimate the relative change in the output of a Monte Carlo simulation based on changes in the input variables. Examples include the simple correlation coefficient, the Spearman rank correlation coefficient and a variety of multiple regression techniques. See Appendix B for a discussion of these techniques.

Monte Carlo methods can also be used to determine the sensitivity over a subset of the output distribution, such as the RME range (i.e., 90^{th} to 99.9^{th} percentiles). For some exposure models, the relative contribution of exposure variables may be different for the high-end exposed individuals than for the entire range of exposures. The general strategy for exploring sensitivity over subsets of risk estimates is to first sort the distribution of simulated output values in ascending (or descending) order, and then apply a sensitivity analysis to the subset of interest (e.g., > 90^{th} percentile).

REFERENCES FOR CHAPTER 2

Cullen, A.C. and Frey, H.C. 1999. Probabilistic Techniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. Plenum Press.

Hamby, D.M. 1994. A Review of Techniques for Parameter Sensitivity Analysis of Environmental Models. *Environ. Monitoring and Assess*. 32: 135-154.

Helsel, D.R. and R.M. Hirsch. 1992. Statistical Methods in Water Resources. Elselvier Science B.V.

Iman, R.L. and J.C. Helton. 1988. An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Anal*. 8: 71-90.

Iman, R.L. and J.C. Helton. 1991. The repeatability of uncertainty and sensitivity analyses for complex probabilistic risk assessments. *Risk Anal*. 11: 591-606.

Morgan, M.G. and M. Henrion. 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press.

Rose, K.A., E.P. Smith, R.H. Gardner, A.L. Brenkert and S.M. Bartell. 1991. Parameter sensitivities, Monte Carlo filtering, and model forecasting under uncertainty. *J. Forecast*. 10: 117-133.

Saltelli, A and J. Marivort. 1990. *Non-Parametric Statistics in Sensitivity analysis for Model Output: A Comparison of Selected Techniques*. Reliability Engineering and System Safety. 28: 299-253.

Shevenell, L. and F.O. Hoffman. 1993. Necessity of uncertainty analyses in risk assessment. *J. Haz. Mat*. 35: 369-385.

Stern, A.H. 1994. Derivation of a target level of lead in soil at residential sites corresponding to a *de minimis* contribution to blood lead concentration. *Risk Anal*. 14: 1049-1056.

U.S. EPA. 1997. *Guiding Principles for Monte Carlo Analysis.* Risk Assessment Forum and National Center for Environmental Assessment. Report No. EPA/630/R-97/001.

CHAPTER 3

SELECTION AND FITTING OF DISTRIBUTIONS

3.0 INTRODUCTION

After completing the sensitivity analysis as described in Chapter 2, the next step is to select the most appropriate distributions to represent the factors that have a strong influence on the risk estimates. This important step in the development of a Monte Carlo model can be very challenging and resource intensive.

EXHIBIT 3-1
STRATEGY FOR SELECTING AND FITTING DISTRIBUTIONS

1. Hypothesize a family of distributions;
2. Estimate distribution parameters; and
3. Assess quality of fit of parameters.

L *Specifying probability distributions for all of the input variables and parameters in a PRA will generally not be necessary.*

If the sensitivity analysis results indicate that a particular input variable does not contribute significantly to the overall variability and uncertainty, then this variable may be represented as a point estimate.

A probability density function (PDF), sometimes referred to as a probability model, characterizes the probability of each value occurring from a range of possible values. One advantage of using a PDF is that it represents a large set of data values in a compact way (Law and Kelton, 1991). For example, a lognormal distribution provides a good fit to a large data set of tap water ingestion rates (n = 5600) among children ages 1 to 11 years (Roseberry and Burmaster, 1992). Therefore, the distribution type (lognormal) and associated parameters (mean and standard deviation) provides a complete characterization of variability in intake rates, from which other statistics of interest can be calculated (e.g., median, 95th percentile, coefficient of skewness, etc.). Reducing a complex exposure model to a series of representative and well-fitting PDFs can facilitate both the quantitative analysis and the communication of the modeling methodology. In PRA, a PDF can also be used to characterize uncertainty. For example, the sample mean ($\bar{x}$) is generally an uncertain estimate of the population mean ($\therefore$) due to measurement error, small sample sizes, and other issues regarding representativeness (see Section 3.3.1). A PDF for uncertainty can be used to represent the distribution of possible values for the true, but unknown parameter. Understanding whether uncertainty or variability is being represented by a PDF is critical to determining how the distribution and parameters are to be specified and used in a PRA.

In general, obtaining representative, high quality data sets to characterize variability and uncertainty in exposure and toxicity can be challenging and resource intensive. Often, more than one probability distribution may appear to be a suitable candidate for characterizing a random variable. Recall from Chapter 2 that even if a variable does not contribute greatly to the variability in risk, it may contribute to the overall uncertainty, and *vice versa*. The choice(s) of distributions and methods of parameter estimation should be discussed with an EPA regional risk assessor prior to initiating a PRA. The workplan for the PRA should clearly document the probability distributions selected for the analysis so that these choices can be reviewed by the Agency. Likewise, any input variable that is eliminated from full probabilistic treatment generally should be identified along with an explanation for its exclusion. This and other important components of the workplan are presented in Chapter 6. Chapter 3 provides guidance on selecting and fitting distributions for variability and parameter uncertainty based on the overall

strategy given in Exhibit 3-1. A discussion on selecting and fitting of distributions is also given in EPA's *Report of the Workshop on Selecting Input Distributions for Probabilistic Assessments* (U.S. EPA, 1999a). The remaining chapters provide guidance on techniques for effectively propagating the distributions through a risk model and communicating the results.

1

EXHIBIT 3-2

DEFINITIONS FOR CHAPTER 3

Bin - Regarding a histogram or frequency distribution, an interval within the range of a random variable for which a count (or percentage) of the observations is made. The number of bins is determined on a case-by-case basis. In general, equal interval widths are used for each bin; however, in some cases (e.g., chi-square test), individual bin widths are calculated so as to divide the distribution into intervals of equal probability.

Goodness-of-Fit (GoF) Test - A method for examining how well (or poorly) a sample of data can be described by a hypothesized probability distribution for the population. Generally involves an hypothesis test in which the null hypothesis H_0 is that a random variable X follows a specific probability distribution F_0. That is, H_0: $F = F_0$ and H_a: $F \ldots F_0$.

Independence - Two events A and B are independent if knowing whether or not A occurs does not change the probability that B occurs. Two random variables X and Y are independent if the joint probability distribution of X and Y factors into the product of the individual marginal probability distributions. That is, f(X, Y) = f(X) A f(Y). Independence of X and Y is *not* synonymous with zero correlation (i.e., Cor(X, Y) = 0). If X and Y are independent, then Cor(X, Y) = 0; however, the converse is not necessarily true (Law and Kelton, 1991) - X and Y may be related in a nonlinear fashion but still maintain Cor(X, Y) = 0.

Independent and Identically Distributed (IID) - Random variables that are independent (see above) and have the same probability distribution of occurrence.

Nonparametric Method - A procedure for making statistical inferences without assuming that the population distribution has any specific form such as normal or lognormal. Sometimes referred to as *distribution free* methods. Common examples are the sign test, Spearman rank correlation, and the bootstrap-t approach.

Parameter - In PRA, a parameter is a constant that characterizes the probability distribution of a random variable. For example, a normal probability distribution may be defined by two parameters (e.g., arithmetic mean and standard deviation).

Parametric Distribution - A theoretical distribution defined by one or more parameters. Examples include the normal distribution, the Poisson distribution, and the beta distribution.

Probability Density Function (PDF) - Same as a probability distribution (see below), although frequently used specifically for a continuous random variable. Probability Mass Function is sometimes used to specifically refer to the probability distribution for a discrete random variable.

Probability Distribution - A table, graph, or formula that associates probabilities with the values taken by a random variable. Also called a *probability model*.

Step Function - A mathematical function that remains constant within each of a series of adjacent intervals but changes in value from one interval to the next. Cumulative distribution functions for discrete random variables are step functions.

Z-score - The value of a normally distributed random variable that has been standardized to have a mean of zero and a standard deviation of one by the transformation Z = (X - :)/F. Statistical tables typically give the area to the left of the z-score value. For example, the area to the left of z = 1.645 is 0.95. Z-scores indicate the direction (+/-) and number of standard deviations away from the mean that a particular datum lies assuming X is normally distributed. Microsoft Excel's *NORMSDIST(z)* function gives the probability p such that p = Pr(Z # z), while the *NORMSINV(p)* function gives the z-score z_p associated with probability p such that p = Pr(Z # z_p).

3.1 CONCEPTUAL APPROACH FOR INCORPORATING A PROBABILITY DISTRIBUTION IN A PRA

A step-wise approach is recommended for incorporating probability distributions in a PRA. Flow charts for specifying distributions for variability and uncertainty are given by Figures 3-1 and 3-2, respectively. Both approaches outline an iterative process that involves three general activities: 1) identifying potentially important sources of variability or uncertainty; 2) evaluating plausible options for distributions and parameter estimates; and 3) reporting the results. These steps are discussed below.

3.2 WHAT DOES THE DISTRIBUTION REPRESENT?

A sensitivity analysis may suggest that characterizing variability or uncertainty in a particular variable in the risk model would be useful for evaluating risk estimates. Various metrics of sensitivity are given in Chapter 2. For Monte Carlo simulations of variability, a distribution is desired that characterizes either inter- or intra-individual variability (depending on the complexity of the model). More specifically, the distribution is intended to be representative of the target population - the receptors that are potentially at risk. The distinction between a target population, a sampled population, and a statistical population should be considered carefully when evaluating information for use in both point estimate and probabilistic risk assessment.

In PRA, there may be a perception that, because a distribution is used rather than a point estimate, there is greater confidence that the available information characterizes the target population. While this may be true in some cases (i.e., more information is available than is reflected by the use of the point estimate), there is always some degree of uncertainty in how representative the available data are, and whether or not the data are sufficient to characterize variability. These issues should generally be considered early in the process of developing a probabilistic model, and are important components of the tiered approach (see Chapter 1, Figure 1-4). The importance of relating the distribution to the target population, clearly distinguishing between variability and uncertainty, and evaluating data representativeness is emphasized below.

3.2.1 CONCEPTS OF POPULATION AND SAMPLING

A target population is the set of all receptors that can potentially be at risk. A risk assessor is interested in knowing features of this population of receptors (e.g., exposure duration, exposure frequency, etc.). The target population is often considered to be the “population of concern”. A sampled population is the set of receptors available for selection and measurement. A sampled population may be the target population or may be a different population that is representative of the target population. A statistical population is an approximation of the target population based on information obtained from the sampled population.

Distributions are generated from representative sample populations to make inferences about the target population. Ideally, a sampled population is a subset of a target population and is selected for measurement to provide accurate and representative information about the exposure factor being studied. However, defining representative samples may be a matter of interpretation.

1

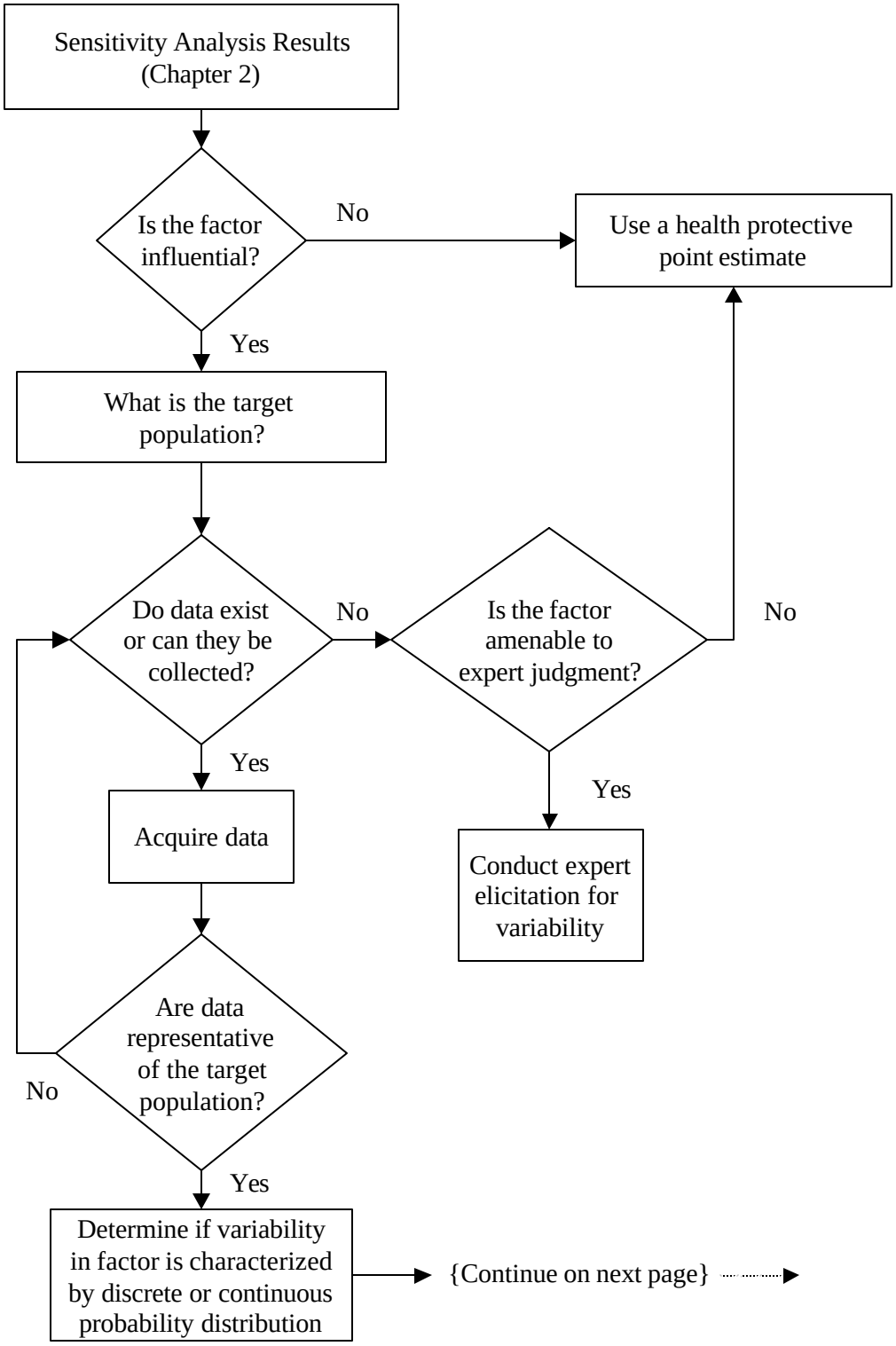


Figure 3-1 (page 1 of 2). Conceptual approach for incorporating probability distributions for variability in PRA.

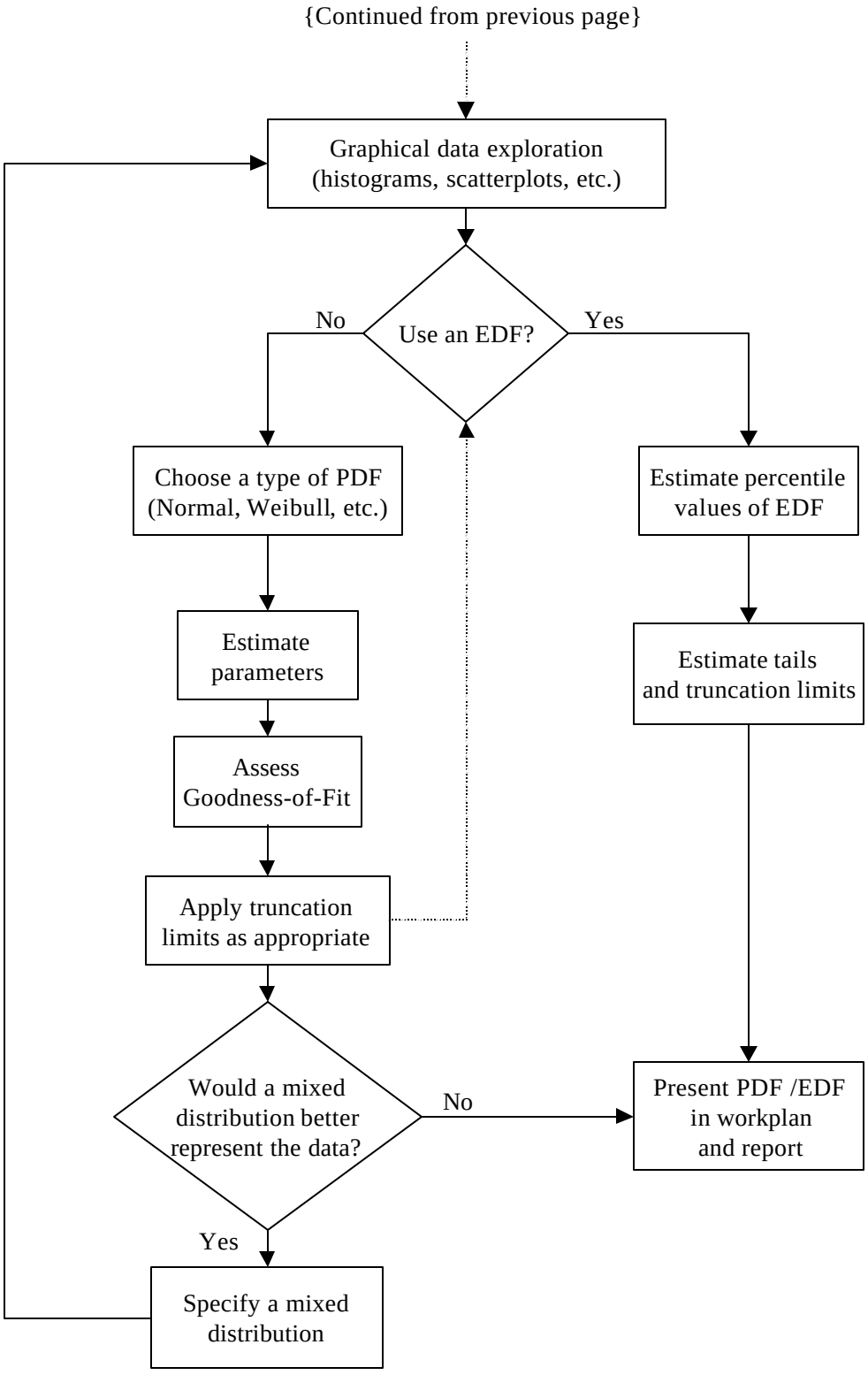


Figure 3-1 (page 2 of 2). Conceptual approach for incorporating probability distributions for variability in PRA.

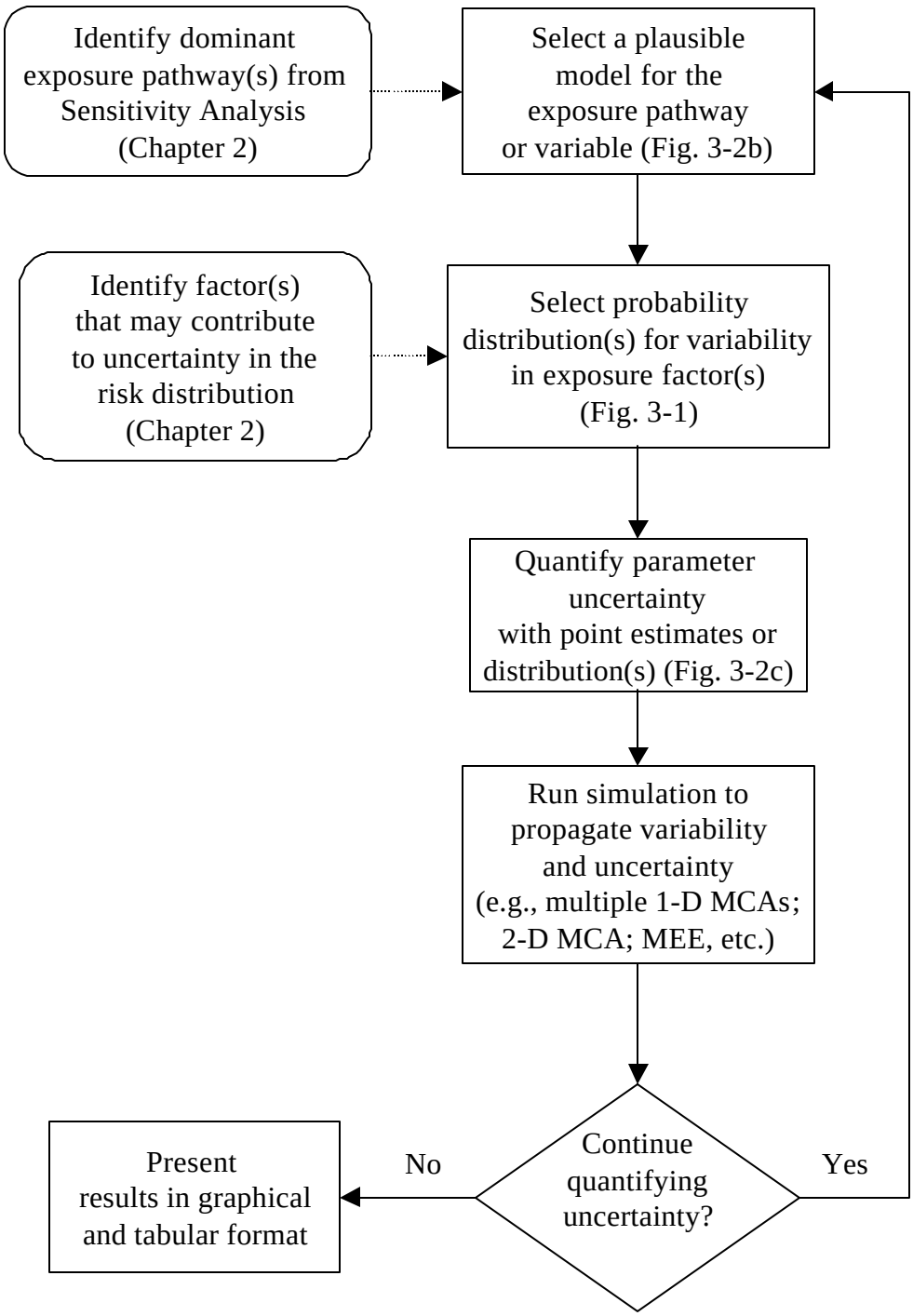


Figure 3-2a (part 1 of 3). Conceptual approach for quantifying model and parameter uncertainty in PRA.

1

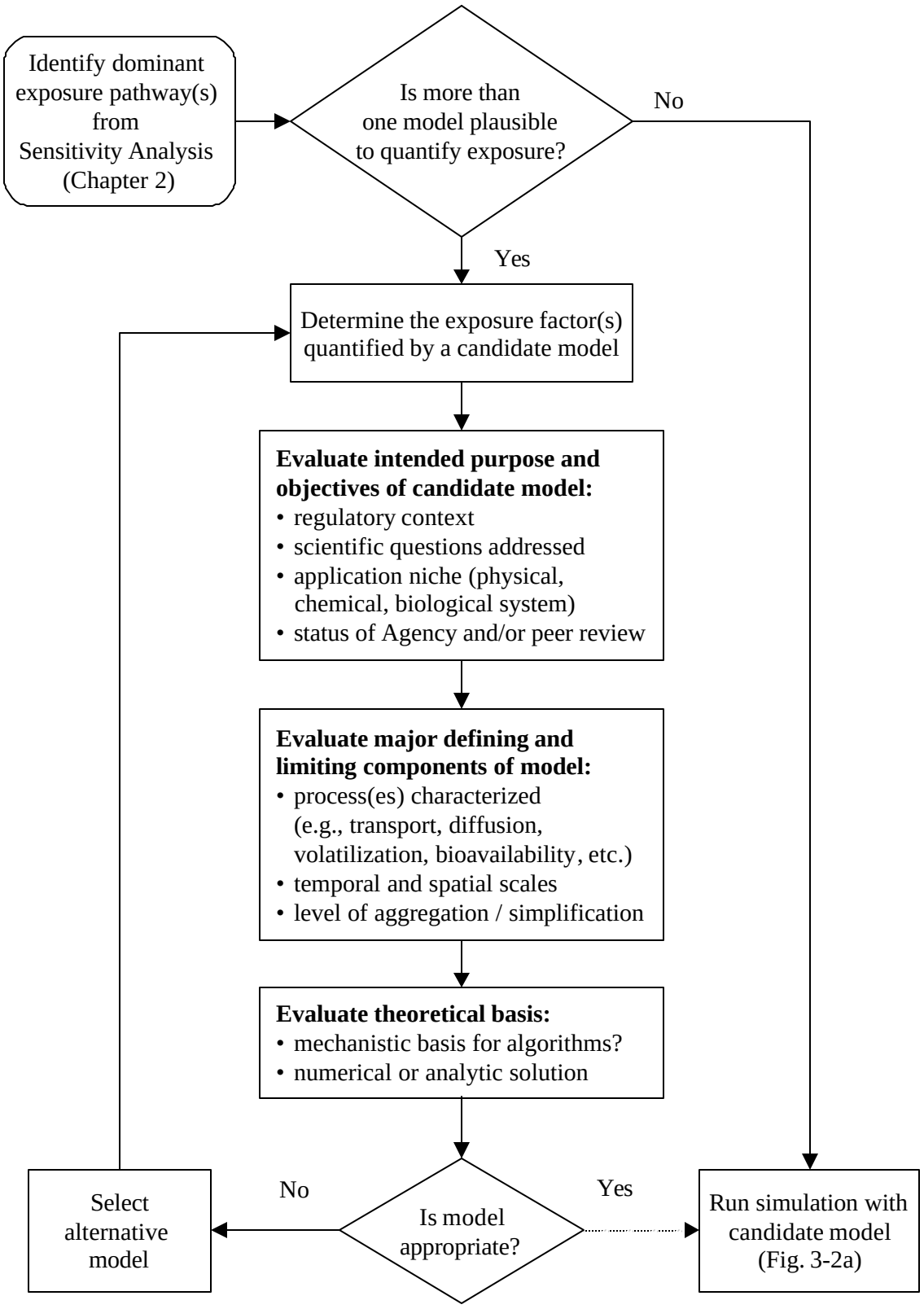


Figure 3-2b (part 2 of 3). Detailed conceptual approach for incorporating model uncertainty in PRA.

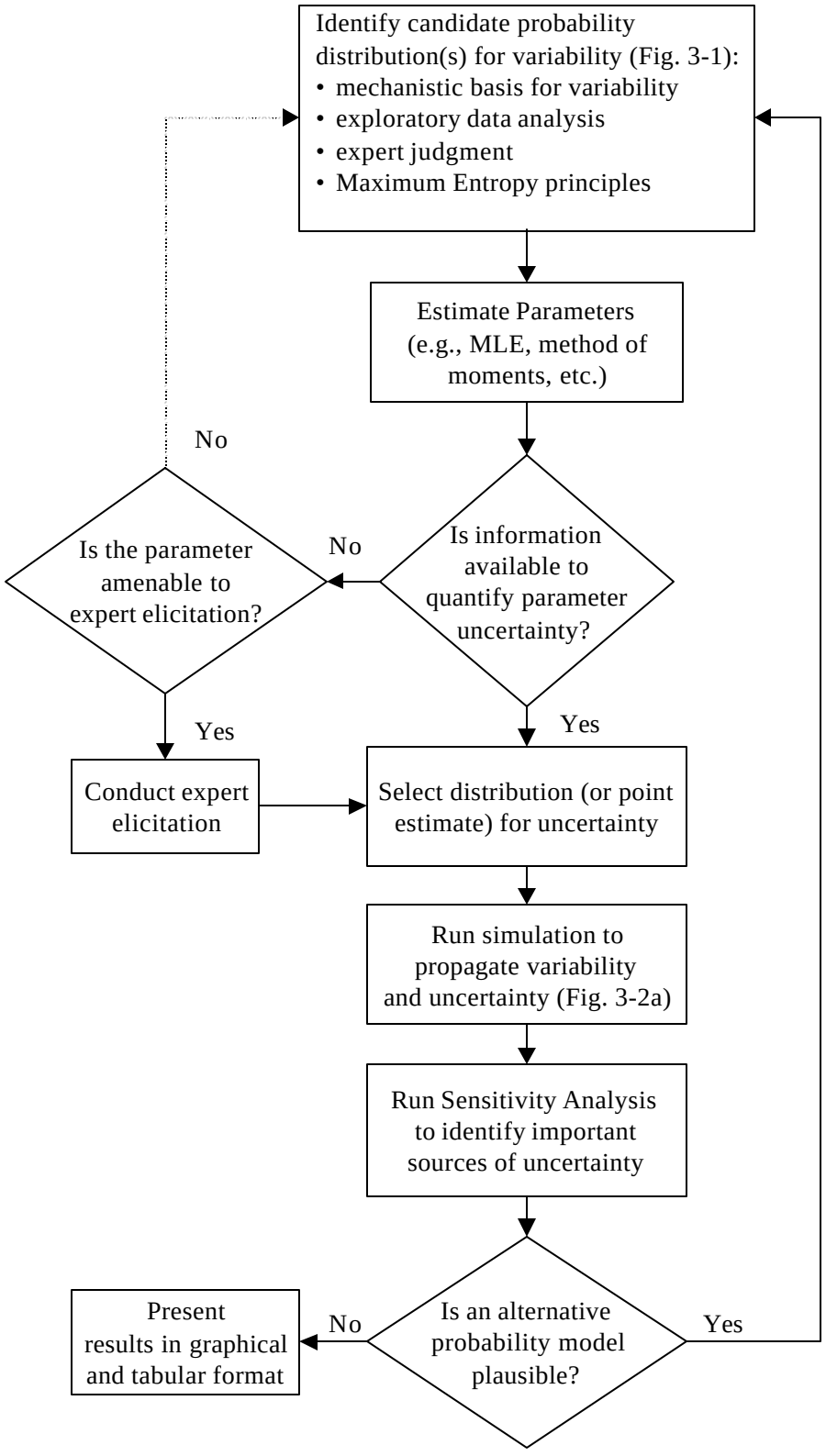


Figure 3-2c (part 3 of 3). Detailed conceptual approach for incorporating parameter uncertainty in PRA.

3.2.2 CONSIDERING VARIABILITY AND UNCERTAINTY IN SELECTING AND FITTING DISTRIBUTIONS

A PDF can represent either variability or uncertainty. Generally, different models will be used to represent variability or uncertainty for an exposure variable. For example, a normal probability distribution may characterize variability in body weight, whereas a uniform distribution may characterize uncertainty in the estimate of the arithmetic mean of the normal distribution. The appropriate interpretation and analysis of data for an exposure variable will depend on whether the distribution will be used to represent variability or uncertainty. Figure 3-1 outlines a process for selecting distributions for variability, whereas Figure 3-2 outlines a process for quantifying both model and parameter uncertainty.

Variability refers to observed differences attributable to true heterogeneity or diversity in a population (U.S. EPA, 1997b). Variability results from natural random processes. Inter-individual variability may stem from environmental, lifestyle, and genetic differences. Examples include human physiological variation (e.g., natural variation in body weight, height, breathing rates, drinking water intake rates), changes in weather, variation in soil types, and differences in contaminant concentrations in the environment. Intra-individual variability may reflect age-specific changes (e.g., body weight and height). Variability is not reducible by further measurement or study. A PDF for variability can be obtained by fitting a distribution to the sample measurements.

Uncertainty refers to the lack of knowledge about specific factors, parameters, or models (U.S. EPA, 1997b). Although uncertainty in exposure and risk assessment is unavoidable due to the necessary simplification of real-world processes, it can be reduced by further measurement and study. Parameter uncertainty may stem from measurement errors, sampling errors, or other systematic errors in the collection and aggregation of data. Model uncertainty may reflect the simplification of a complex process, a mis-specification of the model structure, a misuse or misapplication of a model, and the use of surrogate data or variables. Scenario uncertainty may reflect errors in an exposure model and the use of expert judgment. A conceptual exposure model can be used to provide direction in specifying a probability distribution for uncertainty. For example, the concentration term in a Superfund risk assessment typically represents the long-term average concentration to which a receptor is exposed (see Chapter 4). An uncertainty distribution for the concentration term could be developed from ideas about the statistical uncertainty of estimating the long-term average from a small sample, and the assumption of random movement of the receptors within a defined exposure unit.

This chapter primarily focuses on methods for quantifying uncertainty associated with the selection of a probability distribution, and estimating parameters of a distribution. A probability distribution can be referred to as a type of model in the sense that it is an approximation, and often a simplified representation of variability or uncertainty that combines both data and judgment. A broader use of the term model refers to a representation of a chemical, physical, or biological process. In risk assessment, many different models have been developed, with varying objectives, major defining and limiting components, and theoretical basis. Figure 3-2b provides a general process for exploring model uncertainty of this type. This figure reflects the concepts and spirit of the *Agency Guidance for Conducting External Peer Review of Environmental Regulatory Modeling* (U.S. EPA, 1994). In general, EPA regional risk assessors should be consulted in order to determine the types of models that may be plausible for quantifying exposure at a particular site.

Quantifying parameter uncertainty in a probabilistic model typically requires judgment. When data are uncertain due to, for example, small sample sizes or questionable representativeness (see Section

3.3.1), Monte Carlo simulation can be a useful tool for demonstrating the effect of the uncertainty on the risk estimates. It is most important to model uncertainty when the sensitive input variables (either point estimates or PDFs for variability) are uncertain. While a quantitative uncertainty analysis may complicate a risk management decision by suggesting that risk estimates are highly uncertain, this information can be helpful by focusing additional efforts towards collecting data and reducing uncertainty in the most sensitive input variables. Likewise, if an estimated risk is below a regulatory level of concern, even after quantifying highly uncertain inputs to the model, the risk manager may be more confident in a decision. As emphasized in Figures 3-2a,b, and c, risk assessors should generally refrain from setting *ad hoc* probabilities to different candidate models in a single Monte Carlo simulation. Instead, this guidance strongly recommends exploring model or parameter uncertainty by running a separate simulation with each candidate model. For example, rather than randomly assigning a beta distribution or a lognormal distribution to an exposure variable for each iteration of a simulation, separate simulations should be run with the candidate probability distributions. Similarly, if a range of temporal or spatial scales are plausible for quantifying exposure, multiple simulations should be designed to demonstrate the importance of these assumptions on the risk estimates.

Uncertainty in parameter estimates may be characterized using a variety of methods. Similar to a PDF for variability, a PDF for parameter uncertainty may be represented by a probability distribution with a unique set of parameters. Sometimes the distribution for uncertainty can be specified by knowing (or assuming) a distribution for variability. For example, if X is a normally distributed random variable, the Student's *t* distribution and the chi-square (P^2) distribution can be used as PDFs for uncertainty in the sample mean and variance, respectively. The PDF for both the Student's *t* and chi-square distributions is determined by the sample size (*n*). If a PDF for uncertainty cannot be determined from the PDF for variability, or assumptions regarding the underlying distribution for variability are not supportable, nonparametric or "distribution free" techniques may be used (e.g., bootstrapping). Both parametric and nonparametric techniques may yield confidence intervals for estimates of population parameters.

3.2.3 DEALING WITH CORRELATIONS AMONG VARIABLES OR PARAMETERS

Correlations between exposure variables or between parameters of a probability distribution may also be important components of a probabilistic model. While Monte Carlo simulation software (e.g., Crystal Ball®, @Risk®) is available to incorporate correlations into a model, relevant data for specifying the correlation matrix are generally unavailable. Therefore, in most PRAs, random variables are sampled independently in 1-D MCA and distributions for parameters are sampled independently in 2-D MCA. However, in the absence of data, it may be useful to explore the effect of plausible correlations as part of a more detailed evaluation of uncertainty (i.e., Tier 3 of the PRA process given by Figure 1-4). One approach that can be used to correlate two random variables (or uncertain parameters) is to specify a bivariate normal distribution (Nelsen, 1986; 1987; Brainard and Burmaster, 1992). Example 3-4 in this chapter demonstrates this approach applied to a simple linear regression equation relating contaminant concentrations in soil and dust. In general, because of the complexity of specifying a valid covariance matrix when correlating more than two factors at a time, risk assessors may need to consult a statistician to avoid generating misleading risk estimates.

3.3 DO DATA EXIST TO SELECT DISTRIBUTIONS?

Developing site-specific PDFs for every exposure assumption (or toxicity value, in the case of ecological risk) can be time and resource intensive, and in many cases, may not be worth the effort. For those exposure variables that do exert a significant influence on risk, a PDF may be developed from site-specific data, data sets available in the open literature (e.g., EPA's *Exposure Factors Handbook* [U.S. EPA 1997a]), or from existing PDFs in the literature (e.g., Oregon DEQ, 1998; peer-reviewed publications).

At Superfund sites, perhaps the most common PDF developed from site-specific data will be the media concentration term. The sample (i.e., collection of empirical measurements) will most often be used to estimate either a point estimate of uncertainty (e.g., an upper confidence limit for the arithmetic mean concentration - the 95% UCL), or a distribution that characterizes the full distribution of uncertainty in the mean. Exposure variables such as ingestion rates, exposure duration, and exposure frequency will most likely be derived from existing PDFs or data sets in the open literature. The Agency is currently developing national default PDFs, similar to the *Standard Default Exposure Factors* for point estimates (U.S. EPA, 1991), for a number of the most commonly used exposure pathways and variables (U.S. EPA, 1999b; 1999c). Until these PDFs are developed, PDFs for exposure variables that lack adequate site-specific data will typically be selected from 1) existing PDF; 2) data on the entire U.S. population; or 3) data on subsets of the U.S. population that most closely represent the target population at a site. If risks to a sensitive subpopulation, such as young children, elderly adults, ethnic groups, or subsistence fishermen, are a concern at a site, then existing PDFs or data sets that best characterize these subpopulations would be preferable to national distributions based on the entire U.S. population. If adequate site-specific data are available to characterize any of the exposure variables, distributions can be fit to those data. The mechanics of how to select and fit distributions are discussed in the remainder of this chapter.

An appropriate question to consider when evaluating data sets for use in exposure and risk assessment is, "How many data are enough?". Generally, the larger the sample size (n), the greater one's confidence in the choice of a probability distribution and the corresponding parameter estimates. Conversely, for small n, goodness-of-fit tests (Section 3.5) will often fail to reject any of the hypothesized PDFs. In general, there is no rule of thumb for the minimum sample size needed to specify a distribution for variability or uncertainty. Increasing sample size may be an important risk management consideration when making decisions in the face of uncertainty.

3.3.1 WHAT ARE REPRESENTATIVE DATA?

The question, "What is a representative sample?", is important to address when selecting and fitting distributions to data. Many of the factors that may determine representativeness (e.g., sample size, the method of selecting the target, and sample population (see Section 3.2.1)) are relevant to both point estimate and probabilistic risk assessment. EPA's *Guidance for Data Usability in Risk Assessment* (U.S. EPA, 1992) describes representativeness for risk assessment as the extent to which data define the true risk to human health and the environment.

The goal of representativeness is easy to understand. However, evaluating data to determine if they are representative is more difficult, especially if the problem and decision objectives have not been clearly defined.

The importance of representativeness also varies with the level of complexity of the assessment. If a screening level assessment is desired, for example, to determine if concentrations exceed a health protective exposure level, then representativeness may not be as important as health protectiveness. However, if a complete baseline risk assessment is planned, the risk assessor should generally consider the value added by more complex analyses (e.g., more site-specific data collection, sensitivity analysis, and exposure modeling). A tiered approach for making these decisions for a PRA is presented in Chapter 1; examples of more complex analyses are presented in Appendix E. In addition, the Agency (U.S. EPA, 1999a) summarizes the advantages and weaknesses of proposed checklists for risk assessors to evaluate representativeness of exposure factors data.

A surrogate study is one conducted on a sampled population that is similar to, but not a subset of the target population. When using surrogate data, the risk assessor must generally exercise judgment about the representativeness of the data to the target population. For example, the distribution of body weights of deer mice from two independent samples from similar ecosystems may differ depending on the age structure, proportion of males and females, and the time of year that the samples were obtained. Uncertainties associated with the use of surrogate studies should be discussed in the assessment.

The evaluation of data representativeness will necessarily involve judgment. The workplan should generally include a description of the data, the basis for the selection of each distribution, and the method used to estimate parameters (see Chapter 6). Empirical data (i.e., observations) are typically used to select distributions and derive parameter estimates. One exception is the use of expert judgment or elicitation (see Section 3.3.2) in cases where the quality or quantity of available data are found to be inadequate.

3.3.2 THE ROLE OF EXPERT JUDGMENT

Expert judgment is an inferential opinion of a specialist or group of specialists within an area of their expertise. When there is a data gap for an input variable, expert judgment is appropriate for obtaining distributions. Note that distributions elicited from experts reflect individual or group inferences, rather than empirical evidence. Distributions based on expert judgment can serve as Bayesian priors in a decision-analytic framework and can be modified as new empirical data become available. There is a rich literature regarding the protocol for conducting expert elicitations and using the results to support decisions (Morgan and Henrion, 1990). Elicitation of expert judgment has been used to obtain distributions for use in risk assessments (Morgan and Henrion, 1990; Hora, 1992; U.S. EPA, 1997b) and in developing air quality standards (U.S. EPA, 1982).

Bayesian analysis is a statistical approach that allows the current state of knowledge,

EXHIBIT 3-3

FACTORS TO CONSIDER IN SELECTING A PROBABILITY DISTRIBUTION*

- *Is there any mechanistic basis for choosing a distributional family?*
- *Is the shape of the distribution likely to be dictated by physical or biological properties or other mechanisms?*
- *Is the variable discrete or continuous?*
- *What are the bounds of the variable?*
- *Is the distribution skewed or symmetric?*
- *If the distribution is thought to be skewed, in which direction?*
- *What other aspects of the shape of the distribution are known*?
- *How well do the tails of the distribution represent the observations?*

*Source: U.S. EPA, 1997b

expressed as a probability distribution, to be formally combined with new data to reach an updated information state. In PRA, Bayesian Monte Carlo analysis (Bayesian MCA) can be used to determine the reduction in uncertainty arising from new information. When combined with techniques from decision analysis, Bayesian MCA can help to determine the type and quantity of data that generally should be collected to reduce uncertainty. The benefits and limitations of Bayesian statistics, Bayesian MCA, and decision analysis (i.e., value of information, or VOI), as applied to PRA, are discussed in greater detail in Appendix E.

3.4 FITTING DISTRIBUTIONS TO DATA

Sometimes more than one probability distribution may adequately characterize variability or uncertainty. In general, the preferred choice is the simplest probability model that *adequately* characterizes variability or uncertainty. For example, a log-logistic distribution would not necessarily be selected over a 2-parameter lognormal distribution simply because it was ranked higher in a goodness-of-fit test by a statistical software package. Some distributions (e.g., normal, lognormal) are well known among risk assessors. The statistical properties for these distributions are well understood and the formal descriptions can be quite brief. In addition to the available data, the choice of distributions can be influenced by knowledge of the mechanisms or processes that result in variability.

Important factors to consider in selecting a PDF are described in Exhibit 3-3. An initial step in selecting a distribution is to determine if the random variable is discrete or continuous. Continuous variables take any value over one or more intervals and generally represent measurements (e.g., height, weight, concentration). A mathematical function describes the probability for each value across an interval for a continuous variable. Discrete variables take either a finite or (at most) a *countably infinite* number of values that have only integers. The number of rainfall events in a month is an example of a discrete random variable, whereas the amount of rainfall is a continuous variable. Similarly, the number of fish meals per month is a discrete variable, whereas the average size (mass) of a fish meal is continuous. Unique probabilities are assigned to each value of a discrete variable. Another important consideration is whether there are plausible bounds or limits for a variable. For example, it is highly unlikely that an American adult will weigh less than 30 kg or more than 180 kg. Most exposure variables may assume any non-negative value within a plausible range. A more detailed discussion of factors to consider in selecting a PDF and specifying parameter values is provided below.

3.4.1 CONSIDERING THE UNDERLYING MECHANISM

There may be mechanistic reasons depending on known physical or biological processes that dictate the shape of the distribution. For example, normal distributions result from processes that sum random variables whereas lognormal distributions result from multiplication of random variables. A Poisson distribution is used to characterize the number of independent and randomly distributed events in a unit of time or space. An exponential distribution would describe the inter-arrival times of independent and randomly distributed events occurring at a constant rate. If, instead, the elapsed time until arrival of the k^{th} event is of interest, then the appropriate probability distribution would be the gamma distribution.

- *In all cases, it is incumbent on the risk assessor to explain clearly and fully the reasoning underlying the choice of a distribution for a given exposure variable - primarily from a mechanistic standpoint.*

Appendix C (Table C-1) describes the causal mechanisms or phenomena underlying several commonly used distributions and provides examples of phenomena they describe. Appendix C also illustrates probability distributions (both PDFs and CDFs) commonly used in PRA. While intuitively appealing, identifying a mechanistic basis for a distribution can be difficult for many exposure variables; however, it may be relatively apparent that the variable is bounded by a minimum (e.g., ingestion rate $ 0 mg/day) and a maximum (e.g., absorption fraction # 100%), or that the relevant chance mechanism results in a discrete distribution rather than a continuous distribution, as described above.

3.4.2 EMPIRICAL DISTRIBUTION FUNCTIONS (EDFS)

In some cases, an empirical distribution function (EDF) may be preferred over evaluating the fit of alternative probability models to a data set. EDFs provide a way to use the data itself to define the distribution of the relevant variable. Briefly, an EDF for a random variable is described by a step function based on the frequency distribution of observed values. An EDF for a continuous random variable may be linearized by interpolating between levels of the various bins in a frequency distribution. This CDF for a linearized EDF appears as a line, rather than steps. Example 3-1 illustrates an EDF, linearized EDF, and beta distribution ($\alpha_1 = 0.63$, $\alpha_2 = 2.85$, rescaled to min = 0, max = 364) fit to percentile data for soil ingestion rates in children (Stanek and Calabrese,1995). A plausible range (i.e., minimum and maximum values) was imposed on the data set for this example.

EDFs provide a complete representation of the data with no loss of information. They do not depend on the assumptions associated with estimating parameters for other probability models. EDFs provide direct information about the shape of the distribution, revealing skewness, multimodality, and other features. However, EDFs may not adequately represent the tails of a distribution due to limitations in data acquisition. In the simplest case, an EDF is constrained to the extremes of the data set. This may be an unreasonable restriction if limiting the EDF to the smallest and largest sample values is likely to greatly underestimate the distributional tails. If this is an important source of uncertainty, the risk assessor may choose to extend the tails of the distribution to plausible bounds or to describe the tails with another distribution (see Exhibit 3-4). For example, an exponential distribution may be used to extend the tails based on the last five percent of the data. This method is based on extreme value theory, and the observation that extreme values for many continuous, unbounded distributions follow an exponential distribution (Bratley et al., 1987). As with other probability models, uncertainty in the plausible bounds of an EDF may be reduced by obtaining additional information.

EXHIBIT 3-4

VARIATIONS OF THE EDF*

Linearized - Linearly interpolates between two observations, yielding a linearized cumulative distribution pattern.

Extended - In addition to linearizing (see above), adds lower and upper bounds based on expert judgment.

Mixed Exponential - Adds an exponential upper and/or lower tail to the EDF.

*Source: T. Barry in U.S. EPA, 1999a.

Advantages and disadvantages of using EDFs in PRA are discussed in detail in the *Report of the Workshop on Selecting Input Distributions for Probabilistic Assessments* (U.S. EPA, 1999a).

3.4.3 GRAPHICAL METHODS FOR SELECTING PROBABILITY DISTRIBUTIONS

Graphical methods can provide valuable insights and generally should be used in conjunction with exploratory data analysis. Graphical methods may include frequency distributions, stem-and-leaf plots, dot plots, line plots for discrete distributions, box-and-whisker plots, and scatter plots (Tukey, 1977; Conover, 1980; Morgan and Henrion, 1990).

- *Graphical methods are invaluable for exploring a data set to understand characteristics of the underlying population.*

Together with statistical summaries, graphical data summaries can reveal important characteristics of a random variable, including skewness (asymmetry), number of peaks (multi-modality), behavior in the tails, and data outliers.

Frequency Distribution or Histogram

The frequency distribution, or histogram, is a graphical approximation of the empirical PDF. Frequency distributions can be plotted on both linear and log scales. The general strategy for selecting the number of bins to partition the data is to avoid too much smoothing or too much jaggedness. Equation 3-1 (U.S. EPA, 1999a) provides a starting point for estimating the number of bins based on the sample size (n).

$$Number of\ Bins = 1 + 3.322 \log_{10} n \qquad \text{Equation 3-1}$$

Probability Plotting

Another useful method is probability plotting, also referred to as linear least square regression or regression on ordered statistics. This technique involves finding a probability and data scale that plots the cumulative distribution function (CDF) of a hypothesized distribution as a straight line. The corresponding linearity of the CDF for the sample data provides a measure of the goodness-of-fit of the hypothesized distribution. The general approach involves sorting the sample data in ascending order and converting the ranks to percentiles. The percentile value for the i^{th} rank is calculated according to Gilbert (1987) as:

$$Percentile = 100 \cdot \frac{i - 0.5}{n} \qquad \text{Equation 3-2}$$

An alternative formula is provided by Ott (1995):

$$Percentile = 100 \cdot \frac{i}{n + 1} \qquad \text{Equation 3-3}$$

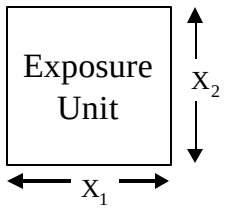
Plotting positions given by Equations 3-2 and 3-3 are special cases of the more general formula given by Equation 3-4 (Helsel and Hirsch, 1992):

$$Percentile = 100 \cdot \frac{i - a}{n + 1 - 2a} \qquad \text{Equation 3-4}$$

where a is a constant that varies from 0 (Equation 3-3) to 0.5 (Equation 3-2).

The percentiles are used to calculate the *z-scores*, which represent the number of standard deviations away from the mean that a particular datum lies assuming the data are normally distributed. For normal distributions, the data are plotted against the z-scores; for lognormal distributions, the data are log-transformed and plotted against the z-scores. In both cases, parameters of the distribution can be estimated from the least-squares regression line. Both Gilbert (1987) and Ott (1995) provide excellent descriptions of the use of probability plotting to derive parameter estimates for a given distribution. Probability plotting techniques with best-fit lines have been used to estimate parameters for a wide variety of distributions, including beta, Weibull, and gamma.

EXHIBIT 3-5
ESTIMATING THE AREA OF A HYPOTHETICAL EXPOSURE UNIT



3.4.4 PARAMETER ESTIMATION METHODS

As a rule, there are often a number of different methods available for estimating a given parameter. The most appropriate method to apply may require judgment, depending on the relative difficulty in applying a method for a particular parameter, as well as the desired statistical properties of the method. The following simple example provides a useful analogy. Suppose that the parameter of interest, A, is the total area of an approximately square exposure unit. If the exposure unit is a perfect square, and the length of and one side (L_1) is known, the area would be equal to L_1^2 (i.e., for a square, $A = L_i^2$). Suppose L is unknown, but two independent measurements, X_1 and X_2, are available to estimate the length (see Exhibit 3-5). If it is assumed that the random variable, L, has a probability distribution with mean ∶, then the area of the square piece of property is $A = ∶^2$. What is a reasonable estimate of the area (i.e., $\hat{A} = \hat{m}^2$) based on X_1 and X_2? Three plausible methods for calculating $\hat{m}^2$ are given below.

1. $\hat{m}_a^2 = \left(\frac{X_1 + X_2}{2}\right)^2$

2. $\hat{m}_b^2 = \frac{X_1^2 + X_2^2}{2}$

3. $m_c^2 = X_1 \cdot X_2$

EXHIBIT 3-6

CRITERIA FOR EVALUATING PARAMETER ESTIMATION METHODS*

Consistency	A consistent estimator converges to the "true" value of the parameter as the number of samples increases.
Efficiency	An efficient estimator has minimal variance in the sampling distribution of the estimate.
Robustness	A robust estimator is one that works well even if there are departures from the assumed underlying distribution.
Sufficiency	An sufficient estimator is one that makes maximum use of information contained in a data set.
Unbiasedness	An unbiased estimator yields an average value of the parameter estimate that is equal to that of the population value.

Because these three estimators will, as a rule, give different answers, it may be useful to set criteria for selecting which one gives the "best" answer. Some of the statistical criteria that are used for this purpose are *consistency*, *efficiency*, *robustness*, *sufficiency*, and *unbiasedness* (see Exhibit 3-6). It turns out, each method is relatively easy to implement, but the third method in the hypothetical example $\hat{m}_c^2$ is preferred because it is a more efficient estimator.

EXHIBIT 3-7

PARAMETER ESTIMATION METHODS

- Method of Matching Moments
- Maximum Likelihood
- Minimum Chi-Square
- Weighted Least-Squares

In many cases, particularly if a model is complex, potential estimators of the unknown parameters are not readily apparent. To assist in developing estimators, several general methods have been developed. Exhibit 3-7 lists some of the more common parameter estimation methods.

Perhaps the simplest method is the method of matching moments (MoMM), also called the method of moments. MoMM is appropriately named, as it involves expressing the unknown parameters in terms of population moments and then "matching", or equating the sample moments to the population moments. For example, the sample mean ($\bar{x}$) and standard deviation (*s*) are estimators for the corresponding population parameters (: and F).

Maximum Likelihood estimation (MLE) is perhaps the most commonly applied method. This is because it can be thought of as an estimate for which the observed data are most "likely". This concept is intuitively appealing. It has also been demonstrated that MLE yields estimators that generally have good properties when evaluated by the criteria listed above. In some cases, these estimators are not *unbiased*; however, this can often be accounted for by "adjusting" the estimator. A familiar example of this adjustment is in estimation of the variance of a normal distribution. The MLE for the variance is biased by a factor of ((*n*-1)/*n*), but this is easily corrected by multiplying the MLE by (*n*/(*n*-1)). For some distributions, calculations of the MLE are straightforward. For example, MLE for parameters of a normal distribution are given by the mean and standard deviation of the sample data, the same as MoMM. MLE for parameters of a lognormal distribution are given by the mean and standard deviation of the log-transformed data, different from MoMM. In general, MLE calculations are complex, and commercial software such as *@Risk*® and *Crystal Ball*® may be used. Also, a more detailed discussion of the derivation and properties of MoMM and MLE can be found in the statistics literature (e.g., Chapter 5 of Mood and Graybill, 1963; Chapter 9 of Mendenhall and Scheaffer, 1973; Section 6.5 of Law and Kelton, 1991; Section 5.6 of Cullen and Frey, 1999).

3.4.5 TRUNCATION

Truncation refers to imposing a minimum and/or maximum value on a probability distribution. The main purpose of truncation is to constrain the sample space to a set of "plausible values". For example, a probability distribution for adult body weight may be truncated at a minimum value of 30 kg and a maximum value of 180 kg in order to avoid the occasional selection of an unlikely value (e.g., 5 or 500 kg). By truncating the tails of a distribution, each risk estimate of a Monte Carlo simulation reflects a combination of plausible input values. The advantage of truncating unbounded probability distributions in PRA is that central tendency and high-end risk estimates will not be biased by unrealistic values. The

disadvantage is that the original parameter estimates of the non-truncated distribution are altered by constraining the sample space. The bias in the parameter estimates increases as the interval between the minimum and maximum truncation limit is reduced. For example, a normal distribution with an arithmetic mean of 100 may be fit to a data set; imposing a truncation limit of 300 may result in a truncated normal distribution with an arithmetic mean of 85. The relationship between the truncated and non-truncated parameter estimates can be determined analytically (Johnson, Kotz, and Balakrishnan, 1995) or approximated using Monte Carlo simulations under both truncated and non-truncated scenarios.

Table 3-1. Theoretical bounds and parameter values for selected distributions.

Probability Distribution	Parameters[1]	Theoretical Bounds
Normal	(:, F)	(-4, + 4)
Lognormal	(:, F)	[0, + 4)
Weibull	("", $)	[0, + 4)
Exponential	($)	[0, + 4)
Gamma	("", $)	[0, + 4)
Beta	("$_1$, "$_2$, a, b)	[a, b]
Uniform	(a, b)	[a, b]
Triangular	(a, m, b)	[a, b]
Empirical (bounded EDF)	(a, b, {x}, {p})	[a, b]

a = minimum; b = maximum; : = mean; F = standard deviation; m = mode;
" = shape parameter; $ = scale parameter; x = value; p = probability.

Truncation is typically considered when using unbounded probability distributions (e.g., normal, lognormal, gamma, Weibull) to characterize variability. Table 3-1 gives the theoretical bounds for selected probability distributions that may be more commonly used in PRA. Truncating the minimum value may also be appropriate for distributions whose minimum is defined as zero (e.g., lognormal, gamma, Weibull). Truncation is generally less important when a PDF is used to characterize uncertainty in a parameter estimate (e.g., arithmetic mean), especially since distributions for uncertainty are often bounded by definition (e.g., triangular, uniform). Bounded continuous distributions, such as the beta distribution or empirical distribution (see Section 3.4.2) are not subject to the parameter bias of truncation, although plausible minimum and maximum values must still be identified.

Identifying appropriate truncation limits that reflect “plausible bounds” for an exposure variable will often require judgment. Given that most data sets represent statistical samples of the target population, it is unlikely that the minimum and maximum observed values represent the true minimum and maximum values for the population. However, there may be physiological or physical factors that can aid in setting plausible truncation limits. For example, the maximum bioavailability of chemicals in the GI tract is 100 percent. Similarly, the solubility of chemicals in aquatic environments (accounting for effects of temperature) will generally be less than the chemical solubility in water free of particulates.

In general, sensitivity analysis can be used to determine if truncation limits are an important source of parameter uncertainty in risk estimates. For exposure variables in the numerator of the risk equation, the maximum truncation limit is of greatest concern. For exposure variables in the denominator of the risk equation, the minimum truncation limit is of greatest concern. Details regarding the fit of the tails of the probability distribution and the effect of truncation on the parameter estimates should generally be included in the workplan.

3.4.6 MAXIMUM ENTROPY APPROACH

In cases where limited information is available for a random variable, a general approach to selecting a distribution for variability or parameter uncertainty may be explored based on the principle of maximum entropy (Jaynes, 1957). Maximum entropy is a technique for determining the distribution that represents the maximum uncertainty allowed by the available information and data (Vose, 1996). Because the shape of the distribution is restricted by the information available, rather than by subjective selection of a distribution, this approach has recently received attention for use in exposure assessment (Lee and Wright, 1994). Although the approach can be used to quickly define distributions that maximize uncertainty, the credibility of the distribution depends on the use of accurate, unbiased information. In general, two or more of the following properties are specified in order to use the approach: minimum, maximum, mean, standard deviation, mode, and percentile value. Table 3-2 summarizes the distribution shapes that represent the maximum entropy for the specified set of constraints. For example, given estimates of the lower bound [min] and arithmetic mean [:] for a random variable, an exponential distribution would be recommended with 8 = 1/ : . Similarly, given estimates of a lower bound [min], upper bound [max] and most likely value [mode] for an arithmetic mean, a triangular distribution would be used to represent parameter uncertainty.

Table 3-2. Maximum entropy inference of distribution shapes corresponding to available information.

Information / Constraints	Distribution Shape
[a, b]	uniform
[a, m, b]	triangular
[a, : , b] or [: , F, a, b]	beta
[: , F]	normal
[a, :] or [a = 0, percentile value]	exponential
[a > 0, percentile value]	gamma

a = minimum; b = maximum; : = mean; F = standard deviation; m = mode

3.5 ASSESSING QUALITY OF THE FIT

The quality of the fit of a distribution refers both to the statistical fit of a theoretical distribution to a data set (i.e., goodness-of-fit tests), and an evaluation of the sensitivity of the risk distribution to assumptions regarding the shape and bounds. Together with graphical exploration (Section 3.4.3), this information may be useful when deciding whether or not to incorporate distributions with particular parameter estimates into a PRA.

3.5.1 WHAT IS A GOODNESS-OF-FIT TEST?

Goodness-of-fit (GoF) tests are formal statistical tests of the hypothesis that the data represent an independent sample from an assumed distribution. These tests involve a comparison between the actual data and the theoretical distribution under consideration.

In statistical hypothesis testing the null hypothesis (H_0) is assumed to be true unless it can be proven otherwise. The "evidence" upon which we base a decision to reject or not to reject H_0 is a random sample. As a general rule H_0 is a "straw man" and the outcome to be demonstrated is stated as the alternative hypothesis (H_a). Typically, we seek to reject H_0 in favor of H_a. For example, with the two sample *t*-test, the null hypothesis is that the means of two populations are equal (not different) and the alternative is that they are different. This is expressed as:

$$H_0: m_1 = m_2$$

$$H_a: m_1 \neq m_2$$

Most often, the hypothesis test is used to show that the means are not equal (i.e., reject H_0 in favor of H_a) in order to state that there is a significant difference between the two populations at a specified significance level (e.g., $\alpha = 0.05$). Thus, the hypothesis test is often referred to as a significance test.

The *p-value* in a statistical test is calculated from a sample and represents the probability of obtaining a value of the test statistic as extreme or more extreme if H_0 is in fact true. When the *p-value* is small it means either that we have witnessed an unusual or rare event (by chance we drew an unusual sample that resulted in the extreme value of the test statistic) or that the null hypothesis is not true. Often a value of 0.05 or 0.01 is designated as a cutoff, or significance level α. If the *p-value* is less than α, (e.g., $p < 0.05$), the null hypothesis is rejected in favor of the alternative, and we state that the test result is statistically significant at level α. This does not mean that we have proven H_a is true. Rather, we are saying that based on our sample results, it is unlikely that H_0 is true.

In a GoF test, the hypothesis test is set up the same way as a "traditional" hypothesis test, but the outcome is viewed a little differently. In GoF tests, we generally seek to *fail* to reject H_0 because the null hypothesis states that the data were obtained from a population described by the specified distribution (F_0). The alternative hypothesis is that the data were obtained from a population described by a different distribution. In most applications of GoF techniques, the alternative hypothesis is composite - it gives little or no information on the distribution of the data, and simply states that H_0 is false (d'Agostino and Stephens, 1986). This can be expressed as:

$$H_0: F = F_0$$

$$H_a: F \neq F_0$$

where F_0 is a specific continuous distribution function, such as the CDF for a normal distribution.

- *Goodness-of-fit tests do not prove that the population could be described by the specified distribution, but rather that this assumption could not be rejected.*

Hence, for example, a *p-value* of 0.05 indicates that the assumption of a specified distribution could be rejected at " = 0.05. Therefore, a larger *p-value* indicates a better fit and provides evidence that the specified distribution may be appropriate. This guidance does not recommend an arbitrary cutoff for the *p-value*. A risk assessor performing a GoF test generally should report the *p-value* and whether the fit is considered "good" or "poor". GoF tests are one tool among several to assess the quality of a distribution.

3.5.2 WHAT ARE SOME COMMON GOODNESS-OF-FIT TECHNIQUES?

The following GoF tests can also be found in most general statistical and spreadsheet software. Both *Crystal Ball*® and *@Risk*® software present the results of chi-Square, K-S, and Anderson-Darling tests in their fitting routines.

Shapiro-Wilk Test

The most widely used GoF test in risk assessment is the Shapiro-Wilk test for normality (Gilbert, 1987). This simple hypothesis test can determine whether or not a small data set ($n < 50$) is normally distributed. The test can also be run on log-transformed data to assess whether the data are lognormally distributed. D'Agostino's test may be used for samples sizes larger than those accommodated by the Shapiro-Wilk test (i.e., $n > 50$) (d'Agostino and Stephens, 1986). In addition, Royston (1982) developed an extension of the Shapiro-Wilk test for n as large as 2000 (Gilbert, 1987).

Probability Plot Correlation Coefficient Test

The correlation coefficient r (or the coefficient of determination , r^2) between the data and the z-scores of a normal probability plot (Filliben, 1975; Helsel and Hirsch, 1992) is a similar to the W statistic of the Shapiro-Wilk test. A detailed comparison of the Shapiro-Wilk test and the product correlation coefficient test is given by Filliben (1975) and d'Agostino and Stephens (1986). Helsel and Hirsch (Table B-3, 1992) summarize critical r* values derived by Looney and Gulledge (1985) for the probability plot correlation coefficient test.

Chi-Square Test

The chi-square test is a general test that may be used to test any distribution (continuous or discrete), and for data that are ordinal (e.g., high / medium / low). Chi-square is a measure of the normalized difference between the square of the observed and expected frequencies. For example, by constructing a frequency distribution of the data with k adjacent bins, $j = 1...k$, the number of data points in the j^{th} bin can be compared with the expected number of data points according to the hypothesized distribution. Note that in the case of continuous, unbounded distributions (e.g., normal), the first and last intervals may include $[-\infty, a_1]$ or $[a_k, +\infty]$ (Law and Kelton, 1991). The chi-square test is very sensitive to the chosen number and interval width of bins - different conclusions can be reached depending on how the intervals are specified. Strategies for selecting bins (e.g., setting interval widths such that there are no fewer than 5 data points expected per bin) are given in the statistical literature (d'Agostino and Stephens, 1986; Law and Kelton, 1991). The test statistic is compared with a value of the chi-square distribution with $\{k - r - 1\}$ degrees of freedom, where r is the number of parameters of the hypothesized distribution (see Table 3-1). As described in Section 3.5.1, in general, higher *p-values* suggest better fits.

Kolmogorov-Smirnov (K-S) Test

The K-S test is a nonparametric test that compares the maximum absolute difference between the step-wise empirical CDF (see Figure 3-3) and the theoretical CDF. Because the maximum discrepancy is compared with the test statistic, K-S is sometimes referred to as a *supremum* test (Cullen and Frey, 1999). In general, lower values of the test statistic indicate a closer fit. The K-S test is most sensitive

around the median of a distribution, and, hence, it is of little use for regulatory purposes when the tails of distributions are most generally of concern (U. S. EPA, 1999a). Although it does not require grouping data into bins like the chi-square test, critical values for the K-S test depend on whether or not the parameters of the hypothesized distribution are estimated from the data set (Gilbert, 1987; Law and Kelton, 1991). The Lilliefors test was developed to surmount this problem when the hypothesized distribution is normal or lognormal (Gilbert, 1987).

Anderson Darling Test

The Anderson-Darling test assesses GoF in the tails (rather than the mid-ranges) of a PDF using a weighted average of the squared differences between the observed cumulative densities. The Anderson-Darling test is sometimes referred to as the *quadratic* test (Cullen and Frey, 1999). The test statistic should be modified based on sample size prior to comparison with the critical value. Like the K-S test, in general, lower values of the test statistic indicate a closer fit (i.e., if the adjusted test statistic is greater than the modified critical value for a specified α, the hypothesized distribution is rejected). The Anderson-Darling test is recommended by this guidance because it places more emphasis on fitting the tails of the distribution.

3.5.3 CAUTIONS REGARDING GOODNESS-OF-FIT TESTS

There are many statistical software programs that will run GoF tests against a long list of candidate distributions. It is tempting to use the computer to make the choice of distribution based on a test statistic. However, GoF tests have low statistical power and often provide acceptable fits to multiple distributions. Thus, GoF tests are better used for rejecting poorly fitting distributions than for ranking good fits. In addition, for many distributions, GoF statistics lack critical values when the parameters are unknown (i.e., estimated from the data). It is most appropriate to form an idea of the candidate distribution based on some well reasoned assumptions about the nature of the process that led to the distribution and then to apply a GoF test to ascertain the fit (U.S. EPA, 1999a). Whenever possible, mechanistic and process considerations should inform the risk assessor's choice of a particular distribution rather than the results of a comparison of GoF tests (Ott, 1995). In addition, the value of graphical evaluations of the fit cannot be overstated.

3.5.4 ACCURACY OF THE TAILS OF THE DISTRIBUTION

The tails of a distribution (e.g., $< 5^{th}$ and $> 95^{th}$ percentiles) for an input variable are often of greatest interest when characterizing variability in risk. Distributions fit to data may not characterize the tails of the distribution in a way that represents the target population. In general, the importance of uncertainty in the fit of the tails of particular distributions should be determined on a site-specific basis. For exposure variables in the numerator of the risk equation, the upper tail is of greatest concern. For exposure variables in the denominator of the risk equation, the lower tail is of greatest concern.

The shape of the input PDFs may have a minimal effect on the risk estimates in the tails of the probability distribution when the mean and variance of the input PDFs are held constant (Finely and Paustenbach, 1994; Hoffman and Hammonds, 1992). In contrast, the tails of the input PDFs generally have a significant influence on the risk distribution.

However, to ensure that the shape of a given PDF does not significantly affect the distribution of possible outcomes, a sensitivity analysis should generally be performed using various plausible shapes of the input PDF. For example, an analysis could be tried two ways substituting a lognormal distribution for

a normal distribution in the description of one of the variables. The sensitivity analysis would quantify the effect of this substitution on a chosen percentile of the output distribution (e.g., 95th percentile).

A common question when developing and evaluating Monte Carlo models is, "How many iterations is enough?". Since Monte Carlo sampling is approximately random, no two simulations will yield the same results (unless the same starting point, or seed, of the random number generator is used). A rule of thumb is that the stability of the output distribution improves with increasing numbers of iterations, although there will always remain some stochastic variability. The stability is generally better at the central tendency region of the output distribution than at the tails; therefore, more iterations may be needed when the risk management decision is associated with the higher percentiles (e.g., > 95th percentile). Risk assessors are encouraged to run multiple simulations (with the same inputs) using different numbers of iterations in order to evaluate the stability of the risk estimate of concern. The results of such an exercise should generally be reported to the Agency when submitting a PRA for review. Note that while the speed of modern computers has essentially eliminated the issue (e.g., 10,000 iterations of most 1-D MCA models can be run in less than 1 minute), it may still be an important issue for more complex modeling approaches such as Microexposure Event analysis and 2-D MCA (see Appendix E).

3.6 PRESENTING INFORMATION ON PROBABILITY DISTRIBUTIONS

Important information for reporting probability distributions in PRA is outlined in Exhibit 3-8.

Graphics developed during the selection and evaluation process (Section 3.4.3) may be included in a PRA report. Such graphs would be particularly useful for communicating potential uncertainties in variables that have the greatest influence on the risk estimates.

In general, each probability distribution used in a Monte Carlo analysis should be presented with sufficient detail that the analysis can be reproduced. This information may be presented in tabular and/or graphical summaries. Important information for a summary table would include a description of the distribution type (e.g., lognormal, gamma, etc.), the parameters that define the distribution, a discussion on how the distribution was selected and the representativeness of the underlying data relative to the exposures being evaluated. If other alternatives were available, the report should generally discuss why a particular selection was made. It should also be made clear if the distribution describes variability or uncertainty. Graphical summaries of the distributions may include both PDFs and/or CDFs, depending on the type of information that is being conveyed (see Chapter 1, Exhibit 1-3). The decision to exclude graphical summaries for some of the more common distributions generally should be made in consultation with the Agency risk assessor.

EXHIBIT 3-8

IMPORTANT INFORMATION FOR REPORTING PROBABILITY DISTRIBUTIONS

- type of distribution and relevant parameter values
- goodness-of-fit approach and statistics if distribution was fit to data
- random sampling technique and number of samples selected
- description of data source, including whether the PDF represents variability or uncertainty
- percentiles of the distribution that are relevant to risk descriptors (i.e., on the risk distribution)

EXAMPLES OF FITTING DISTRIBUTIONS USING GRAPHICAL METHODS, GOODNESS-OF-FIT, AND PARAMETER ESTIMATION

Example 3-1. Empirical Distribution Function (EDF) for Soil Ingestion Rates

This hypothetical example illustrates how graphical methods can be used to select probability distributions for variability based on percentile data reported in the literature. Table 3-3 gives the summary statistics that are reported by Stanek and Calabrese (1995) for average daily soil ingestion rates among young children. Three options are explored for selecting a distribution: 1) empirical distribution function (EDF) represented by a step function; 2) linearized and extended EDF; and 3) continuous parametric distributions (beta and lognormal).

In order to specify an EDF, a plausible range (minimum and maximum) must be inferred using judgment. Exposure factors such as ingestion rate are non-negative variables (i.e., minimum $\geq$ 0); given the relatively low value for the 25th percentile (10 mg/day), it is assumed that 0 mg/day is a reasonable minimum value for this example. If children with pica for soil are excluded from the population of concern, the maximum value may be inferred from the relatively shallow slope at the high-end of the distribution. That is, the 90th percentile is reported as 186 mg/day while the 99th percentile is 225 mg/day, an increase of only 40 mg/day; it is assumed that 300 mg/day is a plausible maximum value for this example. Figure 3-3 illustrates the basic step-wise EDF represented by the reported percentile values, as well as the "linearized, extended EDF" (i.e., linear interpolation between reported values and extended lower and upper tails).

An alternative to relying on a linear interpolation between the percentile values is to fit a continuous probability distribution to the reported percentiles. Since the original data are unavailable, standard goodness-of-fit (GoF) tests for the EDF, such as K-S and Anderson-Darling (d'Agostino and Stephens, 1986), cannot be applied. Note that computer software (e.g., Crystal Ball®, @Risk®) will provide test statistics and corresponding *p-values*, however, these results will (inappropriately) reflect the number of percentile values reported rather than the sample size of the original data. Nevertheless, graphical methods may be employed to assess the adequacy of the fit of various PDFs. In this example, a beta distribution and lognormal distribution were fit to the EDF using Crystal Ball®. Figure 3-4 illustrates the CDFs for both distributions.

The beta distribution appears to more closely match the reported percentile values, especially at the upper tail of the distribution. The lognormal distribution has an unbounded maximum that, for this example, results in an extreme overestimate of the 95th and 99th percentiles. The beta distribution, by definition, is bounded at 0 and 1, and rescaled in this example to a maximum of 364 mg/day.

Table 3-3. CDFs for reported and fitted distributions for ingestion rate (mg/day).

Summary Statistic	Reported Values	Linearized, Extended EDF	Beta Distribution[1]	Lognormal Distribution[2]
minimum	--	0	0	0
25th percentile	10	10	13	11
50th percentile	45	45	44	31
75th percentile	88	88	100	86
90th percentile	186	186	165	216
95th percentile	208	208	205	375
99th percentile	225	225	322	3346
maximum	--	300	364	+ 4

[1] Parameters of best-fit beta distribution: $\alpha_1 = 0.63$, $\alpha_2 = 2.85$, min = 0, max = 364.
[2] Parameters of best-fit lognormal distribution: $\mu = 97.6$, $\sigma = 291.8$.

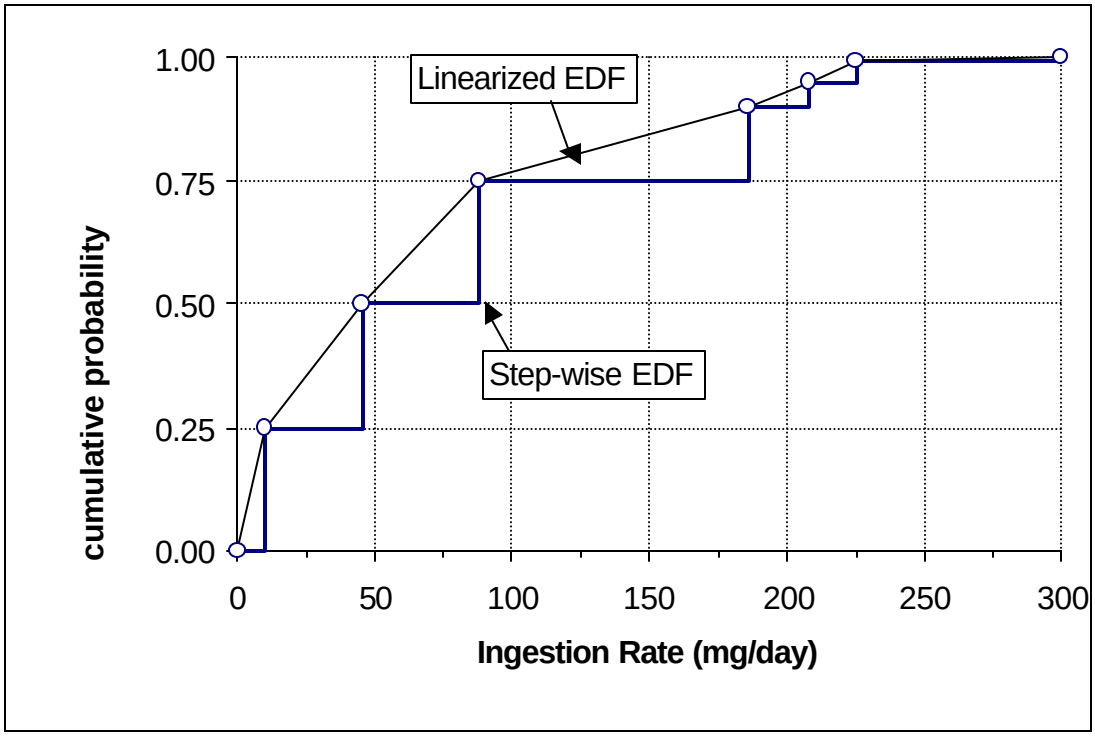


Figure 3-3. Comparison of step-wise EDF and linearized EDF for ingestion rate. The upper and lower tails of both distribution are extended to a plausible range of [0, 300] mg/day.

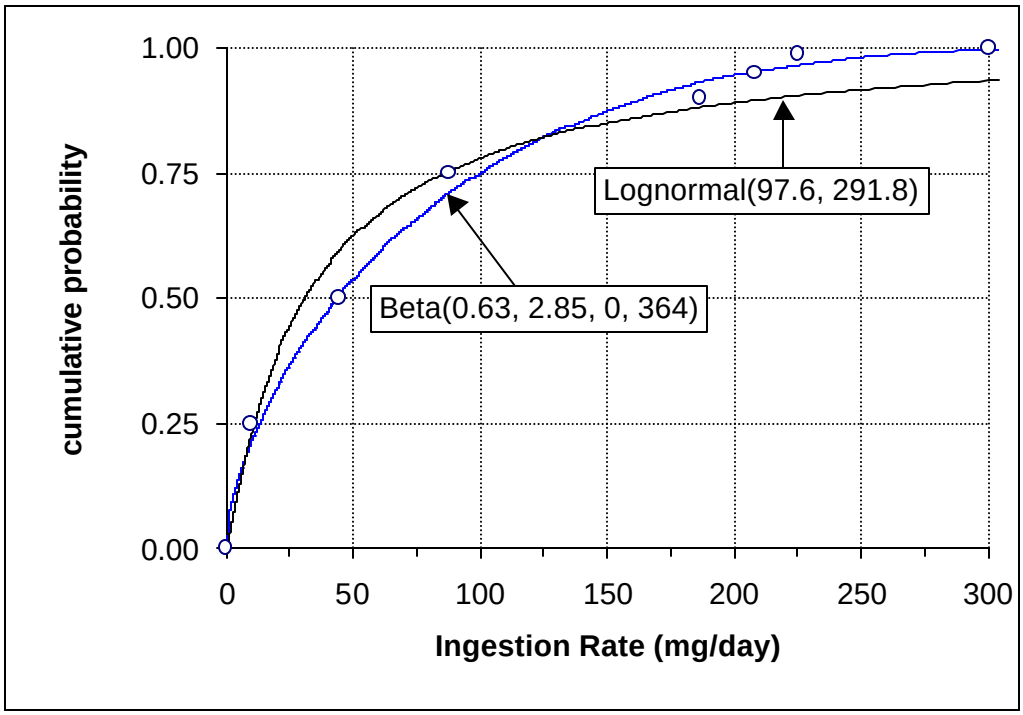


Figure 3-4. Graphical assessment of beta and lognormal distributions fit to the cumulative distribution reported in the literature (circles). The beta distribution provides a closer fit to the percentile values in this example.

Example 3-2. Variability in Lead Concentrations in Quail Breast Tissue

This hypothetical example demonstrates how the combination of graphical methods, GoF tests, and parameter estimation techniques provides strong evidence for selecting and fitting a lognormal distribution. Assume lead concentration in quail is an important variable for a food web model. Site-specific data (n = 62) are used to estimate inter-individual variability in concentration (Table 3-4). The histograms in Figure 3-5 show lead concentrations in quail breast tissue collected near a settling pond at a plating works. Equation 3-1 indicated that 7 bins would be appropriate. The result (top left panel, Figure 3-5) suggests that approximately 80 percent of the values are < 200 ppm and that the probability distribution for variability may be described by a non-negative, long tailed distribution (e.g., exponential, Weibull, lognormal, etc.). However, additional bins are needed to better understand the low-end of the distribution. After increasing the number of bins from 7 to 16 (top right panel, Figure 3-5), graphical evaluation continues to suggest that the distribution is unimodal right skewed. Furthermore, the bottom panel of Figure 3-5 illustrates that increasing the number of bins would not provide better resolution of the low-end of the distribution. For these data, 16 bins appear to provide a reasonable balance between too much smoothing and too much jaggedness.

Probability plots can be used to visually inspect the goodness-of-fit of a specified distribution to the data, and, because the hypothesized distribution yields a straight line, the plots are particularly useful for evaluating deviations at the tails. In addition, parameter estimates can be obtained from the regression lines fit to the data, as discussed below. For this example, two lognormal probability plots are explored to evaluate how well the data can be described by a lognormal distribution (Figure 3-6). The top panel gives the *z-score* on the abscissa (the "x" axis) and ln[concentration] on the ordinate (the "y" axis), while the bottom panel gives ln[concentration] on the abscissa and *z-score* on the ordinate. Plotting positions for both methods were calculated using Equation 3-2. Equally plausible parameter estimates can be obtained from regression lines using either plotting method; however, the approach shown in the top panel may be easier to implement and interpret.

Despite the relatively large sample size of n = 62, GoF tests generally fail to reject lognormality (i.e., normality of the log-transformed data) in this example. For the probability plot correlation coefficient test (Filliben, 1975; Looney and Gulledge, 1985), if r < r* at a specified α, normality is rejected. For this example, r is 0.988, and r* is between 0.988 and 0.989 for n = 62 and α = 0.25; therefore, the *p-value* for the concentrations is approximately 0.25 and one fails to reject normality at $\alpha \leq 0.25$. D'Agostino's test yields essentially the same conclusion, with a calculated Y value of -1.9166. For this data set, with n = 62 and α = 0.10, one rejects normality if Y < -2.17 or Y > 0.997 (see Table 9.7 in d'Agostino and Stephens, 1986); therefore, since Y is within this interval, one fails to reject the normal distribution. However, for α = 0.20, the rejection criteria is [Y < -1.64 or Y > 0.812], Y falls outside the low-end of the interval, resulting in a rejection of the normal distribution. For this data set, the *p-value* associated with d'Agostino's test is slightly less than 0.20 and one fails to reject normality at $\alpha < 0.20$.

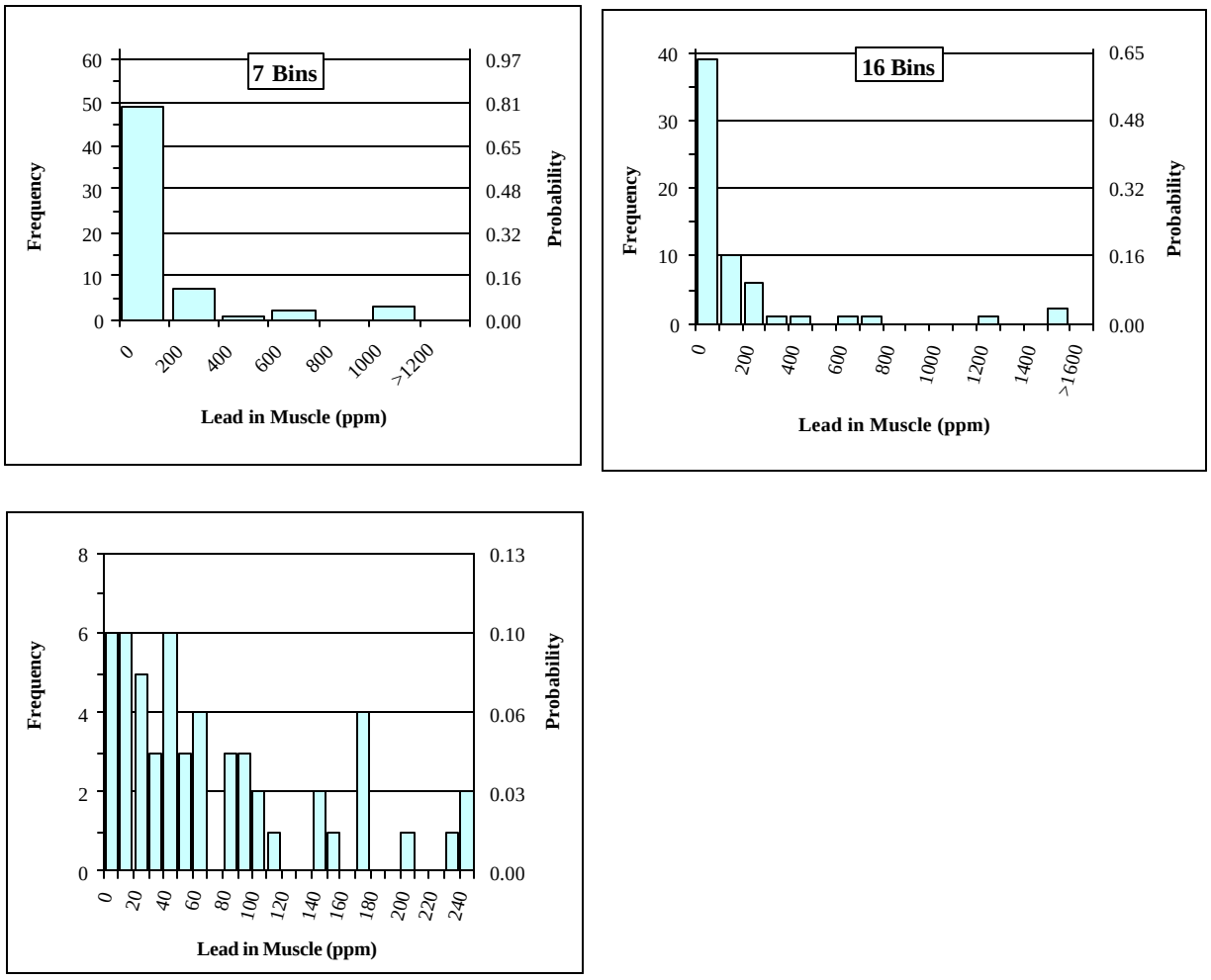


Figure 3-5. Histograms of lead concentrations in quail breast muscle (n = 62). The top left panel shows the result with seven bins; the top right panel shows the result with sixteen bins, the bottom panel uses bin widths of 10 ppm to highlight the lower tail (< 250 ppm) of the distribution.

Table 3-4. Sample values of lead concentration (ppm) in quail breast muscle (n = 62).

0.45	15.8	36.6	57	91	173	265
2.1	16	40	59.6	94.2	175.6	322
5.4	16.7	40.1	61.4	99	176	490
7.8	21	42.8	62	107	177	663.4
7.8	23	44	64	109	205	703
8.8	24	46	64	111	239	1231
11.8	24.8	47	84.6	149	241	1609
12	29.2	49	86.6	149	245	1634
15	35.5	53	86.8	154	264	

Different methods for obtaining the parameter estimates for the lognormal distribution can be explored in this example. For the lognormal distribution, Maximum Likelihood Estimate (MLE) and Method of Matching Moments (MoMM) simply requires calculating the mean and standard deviation of the log-transformed sample data. For the lognormal probability plot method, the parameters can be obtained directly from the least squares regression line expressed as follows:

$$\ln(x) = [slope]z + [intercept] \qquad \text{Equation 3-5}$$

such that exponentiating the intercept will give the geometric mean (GM) and exponentiating the slope will give the geometric standard deviation (GSD) (see footnote 3 of Table 3-5). Both the MLE and MoMM estimates will generally match the arithmetic mean of the log-transformed data (i.e., intercept) determined from lognormal probability plots; however, estimates of the standard deviation (i.e., slope) will vary (Cullen and Frey, 1999). In general, the probability plot method yields estimates of the standard deviation that are less than or equal to that of MoMM and MLE, and the results yield closer estimates as the correlation coefficient of the probability plot increases (Cullen and Frey, 1999). Table 3-5 summarizes the parameter estimates using MLE, MoMM, and the two lognormal probability plotting techniques described above. The corresponding parameter estimates for the untransformed data are also presented.

In this example, the strong linearity of the probability plots ($r^2 = 0.98$) shown in Figure 3-6 is an indication that a lognormal distribution is a reasonable model for describing variability in concentrations. The tails of the distributions fit the data fairly well, although the bottom panel suggests that the lognormal distribution slightly overestimates the lower tail. Furthermore, the parameter estimates of the lognormal distribution using probability plotting closely match the estimates using MLE and MoMM.

$$geometric\ mean = \exp[m^*]$$

$$geometric\ standard\ deviation = \exp[s^*]$$

$$arithmetic\ mean = \exp[m^* + 0.5s^{*2}]$$

$$standard\ deviation = \exp[m^*]\left(\exp[s^{*2}]\exp[s^{*2} - 1]\right)^{0.5}$$

Table 3-5. Parameter estimates for lognormal distribution of lead concentrations (ppm).

Parameter Estimation Method	Log-transformed Data		Untransformed Data[3]	
	mean [$\hat{m}$]	stdev [$\hat{s}$]	mean [$\hat{m}$]	stdev [$\hat{s}$]
Maximum Likelihood Estimate (MLE)	4.175	1.522	207	626
Method of Matching Moments (MoMM)	4.175	1.522	207	626
Log Probability Plot[1]	4.175	1.507	203	597
Log Probability Plot[2]	4.175	1.543	214	670

[1] Least squares regression line for Figure 3-6, top panel.

[2] Least squares regression line for Figure 3-6, bottom panel.

[3] For a lognormal distribution, the following equations can be used to convert parameters of the normal distribution of log-transformed data to corresponding parameters of the lognormal distribution of untransformed data. Assume :* and F* are the arithmetic mean and standard deviation, respectively, for the normal distribution of log-transformed data.

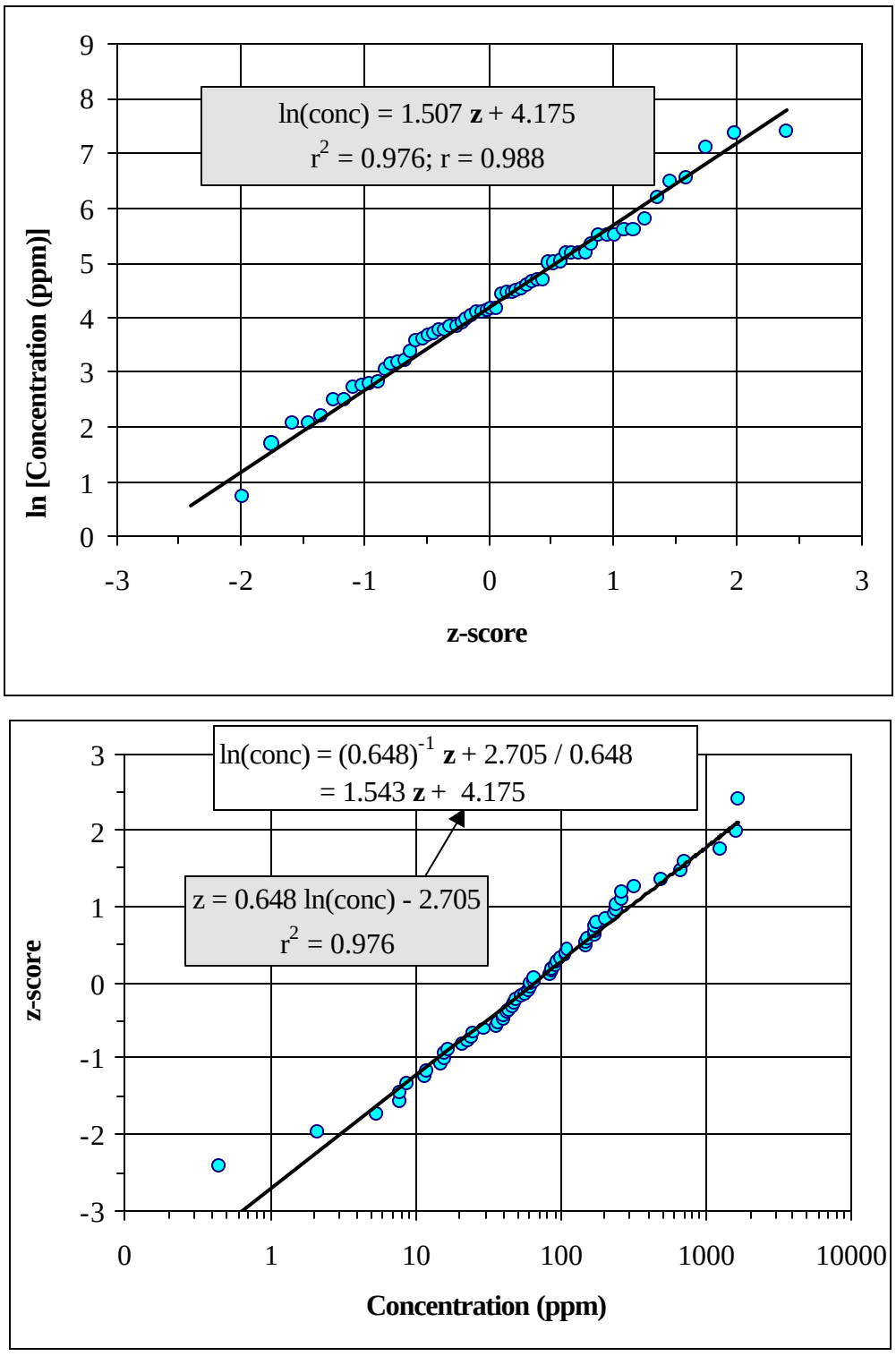


Figure 3-6. Lognormal probability plots of lead in mourning dove breast tissue. Top panel gives ***z*** on the abscissa and ln[concentration] on the ordinate. Bottom panel gives ln[concentration] on the abscissa and ***z*** on the ordinate. Equally plausible parameter estimates can be obtained from regression lines using either plotting method. Bottom panel requires an additional step to express the equation that yields parameter estimates [ln(x) = (slope) z + (y-intercept)], where the slope estimates the standard deviation of ln(x) and the y-intercept (at z = 0) estimates the arithmetic mean of ln(x).

Example 3-3. Variability in Meal Sizes Among Consuming Anglers

A creel survey of anglers consuming contaminated fish was performed to estimate variability in fish meal sizes. The anglers were asked how many people would eat their fish. The lengths of the fish were measured and a regression equation was used to calculate the corresponding weights. The portion of the fish mass that is consumed was assumed to be 40% (e.g., fillets). Results given in Table 3-6 are expressed in units of grams of fish per meal.

The appearance of the histograms (Figure 3-7) suggests that the sample (n = 52) may have been selected from more than one distribution.

A normal probability plot of the meal sizes (Figure 3-8) shows a departure from linearity. Specifically, there appears to be a "kink" in the probability plot at about 400 g/meal, suggesting that the sample may have been obtained from two unique distributions. Both the Filliben test and Shapiro-Wilk test indicated a significant departure from normality at " = 0.01. Parameters may be read directly from the equations of the regression lines on the right hand panel of the graph. MoMM and MLE gave similar estimates.

Table 3-6. Meal size (g/meal) (n = 52).

65	182	310	405
74	208	314	415
74	221	318	416
77	226	318	477
90	241	327	531
110	248	332	572
111	253	336	608
133	260	337	745
143	261	350	831
150	281	351	907
163	303	360	1053
163	305	365	1189
174	305	390	1208

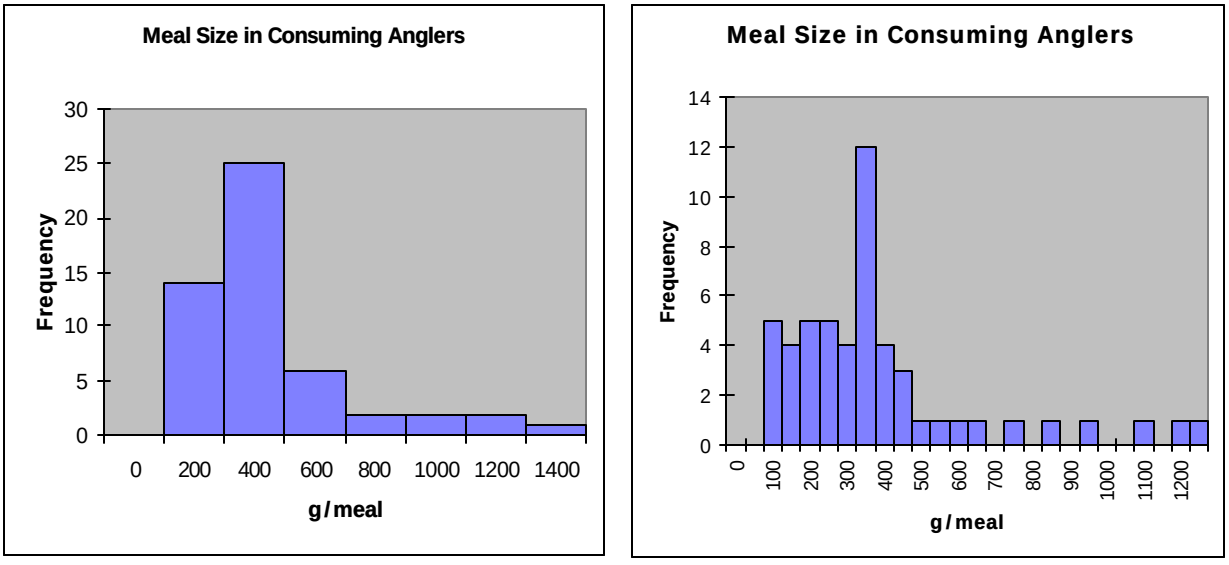


Figure 3-7. Histograms of meal size (n = 52) among consuming anglers. Left panel uses 7 bins, while the right panel uses 14 bins.

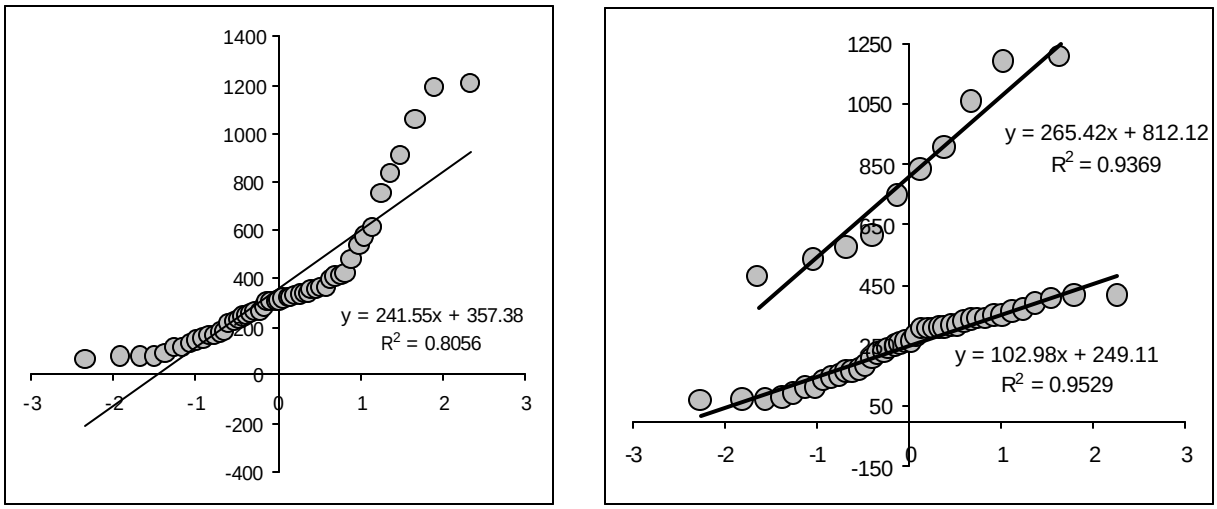


Figure 3-8. Probability plot of meal size data from consuming anglers. The left panel shows the combined data, with a departure from linearity at ~ 400 g/meal. The right panel shows the data split between high consumers (top line) and low consumers (bottom line); note that separate lognormal probability plots were reconstructed for both subsets of the data.

Example 3-4. Bivariate Normal Distributions

One approach that can be used to correlate two random variables is to specify a bivariate normal distribution. A brief explanation of the bivariate distribution is presented followed by an example comparing assumptions of no correlation and perfect correlation.

THIS EXAMPLE PRESENTS….

1. description of the assumptions associated with the bivariate normal distribution;
2. guidance on simulating the bivariate normal distribution for two random variables; and
3. application of bivariate normal to a simple linear regression equation relating contaminant concentrations in soil and dust (see Figure 3-9). Results are compared to the assumption of no correlation and perfect correlation.

Conditional Normal Distributions

Suppose that two random variables, *X* and *Y*, jointly follow a bivariate normal distribution with means μ_X and μ_Y, variances σ^2_X and σ^2_Y, and correlation coefficient ρ. Then the marginal distribution of *X* is normal with mean μ_X and variance σ^2_X, and the marginal distribution of *Y* is normal with mean μ_Y and variance σ^2_Y[1].

Assume we are interested in the conditional distribution of *X* given a certain value for *Y*. For example, if *X* and *Y* are positively correlated, we would expect that relatively high values of *X* tend to correspond with relatively high values of *Y*. The conditional distribution of *X* given that *Y* = y, where y represents a specific value for the random variable *Y*, is a normal distribution with:
Note that the mean of the conditional distribution of *X* is a function of the given value of *Y* but the variance depends only on the degree of correlation.

$$mean = m_X + r\frac{s_X}{s_Y}(y - m_Y), \quad and$$
$$variance = s_X^2(1 - r^2)$$

Equation 3-6

Likewise, the conditional distribution of *Y* given that *X* = x, is also normal with:

$$mean = m_Y + r\frac{s_Y}{s_X}(x - m_X), \quad and$$
$$variance = s_Y^2(1 - r^2)$$

Equation 3-7

These general equations can be used to generate a correlated pair (*X*, *Y*), as described below.

[1] Note, however, that this does not necessarily go the other way. If we start with the assumption that *X* and *Y* each have normal distributions, then it is not necessarily the case that the joint distribution of *X* and *Y* is a bivariate normal.

General Approach for Correlating X and Y

To generate a correlated pair (X, Y), first generate X using a random value Z_1 from the standard normal distribution:

$$X = m_X + s_X \cdot Z_1 \qquad \text{Equation 3-8}$$

Next, express Y as a function of the conditional mean and variance of Y given X and a second standard normal variate Z_2:

$$Y = m_Y + s_Y \cdot Z_2 \qquad \text{Equation 3-9}$$

and generate a correlated Y by plugging Equation 3-7 into Equation 3-9. Using algebra, the combined equations yield the following simplified expression for generating Y:

$$Y = m_Y + s_Y \left[(r \cdot Z_1) + \sqrt{1 - r^2} \cdot Z_2 \right] \qquad \text{Equation 3-10}$$

The important component of this equation is that two random variates are needed (Z_1 and Z_2).

Special Case Approach for Perfect Correlation between X and Y (ρ = 1.0)

An alternative, but less general approach would be to obtain Y by first generating a normal variate X (Equation 3-8) and then plugging that value into the regression equation of Y on X to obtain the associated value of Y. The regression of Y on X for a bivariate normal is given by Equation 3-11:

$$Y = m_Y + r \frac{s_Y}{s_X} (X - m_X) \qquad \text{Equation 3-11}$$

Plugging Equation 3-8 into Equation 3-11, yields the following simplified expression for generating Y:

$$Y = m_Y + r \cdot s_Y \cdot Z_1 \qquad \text{Equation 3-12}$$

While Equation 3-11 maintains a correlation between X and Y, it will underestimate parameter uncertainty by the factor [σ_Y (1 - ρ^2) Z_2]. Equations 3-10 and 3-12 are equal only for the special case of perfect correlation (ρ = 1.0) between X and Y. Therefore, the more general bivariate normal distribution approach (given by Equations 3-8 to 3-10) is recommended for correctly correlating X and Y because it provides a more robust estimate of parameter uncertainty.

Example: Simulation of a Bivariate Normal Distribution for Parameter Uncertainty

Suppose that we wish to simulate uncertainty in the relationship between zinc concentrations in soil (ZnS) and dust (ZnD). For simplicity, assume a simple linear regression equation is selected as an appropriate model to relate the two exposure variables:

$$ZnD = b_0 + b_1 \cdot ZnS \qquad \text{Equation 3-13}$$

where β_0 and β_1 are the intercept and slope parameters that describe the regression line. For this example, β_0 represents X and β_1 represents Y in Equations 3-6 to 3-12. Let b_0 and b_1 be the least squares estimates of β_0 and β_1, respectively. Then, under the usual assumptions regarding the regression model, b_0 and b_1 have a joint normal distribution with the following parameters:

$$\mu_{b0} = b_0$$

$$\mu_{b1} = b_1$$

$$\sigma^2_{b_0} = \sigma^2 \frac{\sum X_i^2}{n \sum (X_i - \bar{X})^2}$$

$$\sigma^2_{b_1} = \frac{\sigma^2}{\sum (X_i - \bar{X})^2}$$

$$Cov(b_0, b_1) = \frac{-\bar{X} \cdot \sigma^2}{\sum (X_i - \bar{X})^2}$$

$$\rho = \frac{Cov(b_0, b_1)}{\sigma_{b_0} \cdot \sigma_{b_1}}$$

where X_i is the i^{th} observation of ZnS and Y_i is the i^{th} observation of ZnD. In addition, the population variance σ^2 is estimated by the sample variance s^2:

$$s^2 = \frac{\sum (Y_i - (b_0 + b_1 \cdot X_i))^2}{n - 2} \qquad \text{Equation 3-14}$$

Application of Bivariate Normal Distribution to Correlate Concentrations of Zinc in Soil and Dust

Assume random sampling of soil and dust zinc concentrations (ZnS, ZnD) occurs in a residential area. Composite samples of soil and dust are collected from 21 locations such that samples are paired (i.e., each soil sample is co-located with a dust sample) (Table 3-7). First the relationship between the ZnS and ZnD is evaluated using simple least-squares regression. Next, the bivariate normal distribution for the slope (β_1) and intercept (β_0) is determined, yielding an arithmetic mean and standard deviation for each parameter (μ_{b0}, $\sigma^2_{b0,}$ μ_{b1}, and σ^2_{b1}), and correlation coefficient ρ between β_1 and β_0. In this context, the bivariate normal distribution may be considered a distribution for uncertainty in the parameter estimates.

Three simulation methods are employed to demonstrate the effect of assuming a bivariate normal distribution for parameters vs. perfect correlation, or independent parameters. Specifically:

1. The slope and intercept of the regression line are described by a specific form of the bivariate normal distribution (i.e., follow *Steps 1, 2* in Exhibit 3-9, and use Equation 3-10 instead of *Step 4*);

2. The slope and intercept of the regression line are described by a general form of the bivariate normal distribution (i.e., follow *Steps 1-4* in Exhibit 3-9); and

3. The slope and intercept of the regression line are described by independent normal distributions (i.e., follow *Steps 1-4* in Exhibit 3-9, but omit the correlation coefficient ρ in *Steps 2 and 4*).

EXHIBIT 3-9

STEPS FOR SIMULATING UNCERTAINTY IN LINEAR REGRESSION EQUATION USING A BIVARIATE NORMAL DISTRIBUTION TO CORRELATE PARAMETERS (θ_0, θ_1)

1. Select Z_1 from a standard normal distribution $Z \sim N(0, 1)$;
2. Calculate β_0 using Equation 3-8, where $X = \beta_0$, $\sigma_x = \sigma_{b0}$, and $\sigma^2_x = \sigma^2_{b0}$;
3. Select Z_2 from a standard normal distribution $Z \sim N(0, 1)$; and
4. Calculate β_1 using Equation 3-10, where $Y = \beta_1$, $\sigma_y = \sigma_{b1}$, $\sigma^2_y = \sigma^2_{b1}$, ρ = correlation between β_0 and β_1.

Fore each approach, Monte Carlo simulations with i = 5000 iterations were run to determine the set of parameter values (β_0, β_1) for a simple linear regression equation. Typically, the uncertainty in the parameter estimates is not accounted for when simple linear regression equations are used to related two exposure variables in a model. Such an approach may fail to account for important sources of parameter uncertainty. Figure 3-10 (middle panel) illustrates the preferred approach for characterizing parameter uncertainty based on the bivariate normal distribution. {Note that the correlation coefficient relating the intercepts and slopes generated from the simulation is consistent with the correlation coefficient that describes the bivariate normal distribution; this is a good check that the simulation was set up correctly and run for a sufficient number of iterations.} These results are contrasted with results using a form of the bivariate normal (Equation 3-10) that underestimates uncertainty (top panel) unless parameters are perfectly correlated. In addition, the simplistic approach of sampling from independent normal distributions (bottom panel), yields a "shot gun" scatterplot. Sampling from independent normal distributions results in unlikely extreme combinations of the slope and intercept more often than the correct bivariate normal approach; propagating this bias through a risk model may severely bias estimates of uncertainty in risk.

Table 3-7. Zinc concentrations in paired (i.e., co-located) soil and dust samples (ppm) for n = 21 locations.

Sample	Soil (X_i)	Dust (Y_i)	Sample	Soil (X_i)	Dust (Y_i)
1	120	216	12	560	200
2	190	149	13	560	256
3	270	83	14	720	496
4	285	508	15	800	239
5	310	215	16	880	203
6	340	219	17	910	757
7	350	203	18	1035	676
8	380	101	19	1445	426
9	440	178	20	1600	522
10	480	232	21	1800	276
11	560	199			

Bivariate Normal Distribution for Parameters of the Regression Equation	
B_0, mean	173.9
B_0, variance	4162.2
B_1, mean	0.193
B_1, variance	0.0063
s^2	27857.4
Cov (B_0, B_1)	-4.2428
r	-0.8254

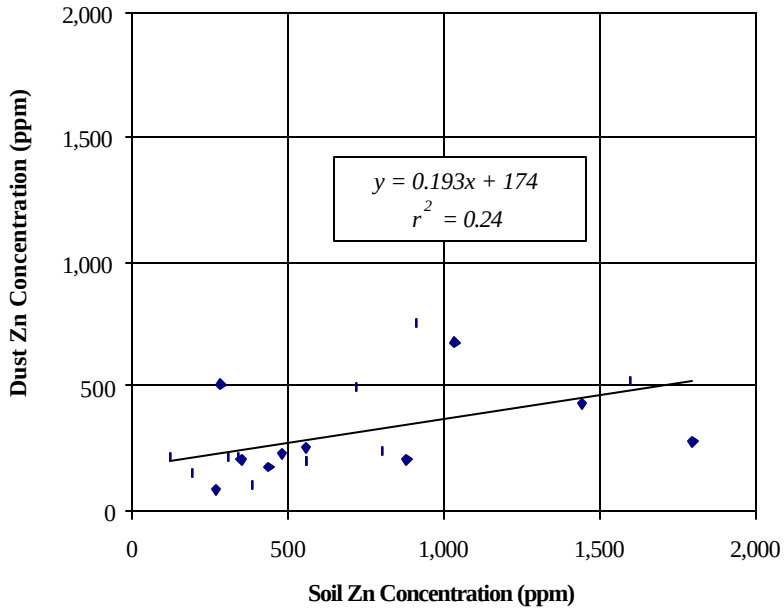


Figure 3-9. Simple linear regression of zinc concentrations in

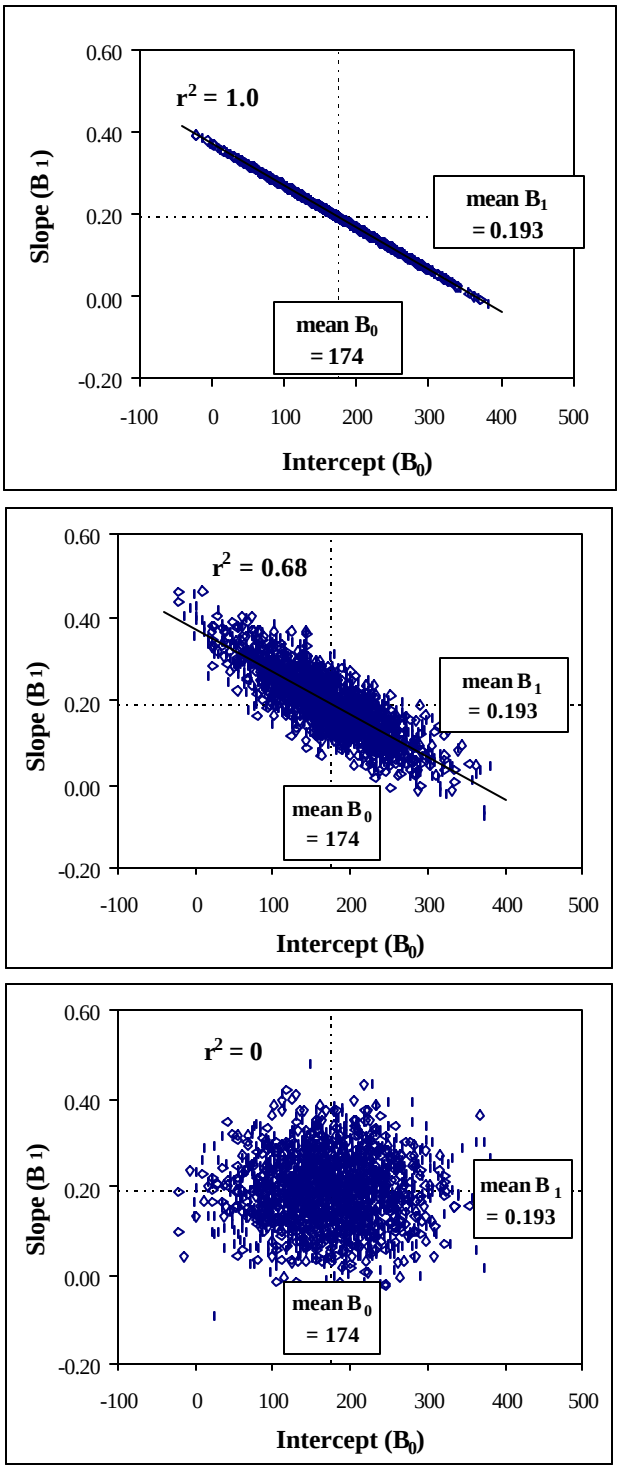


Figure 3-10. Results of Monte Carlo simulation (n = 5000 iterations) to estimate the slope and intercept of a regression equation. Top panel reflects the bivariate normal distribution for the special case that fails to capture the parameter uncertainty; middle panel reflects the preferred bivariate normal distribution with ρ = -0.825 based on empirical paired data; bottom panel reflects sampling from independent normal distributions.

REFERENCES FOR CHAPTER 3

Barry, T. 1999a. Personal Communication.

Brainard, J. and D.E. Burmaster. 1992. Bivariate distributions for height and weight of men and women in the United States. *Risk Anal.* 12(2): 267-275.

Brately, P., B.L. Fox and L.E. Schrage. 1987. *A Guide to Simulation.* Springer-Verlag. New York.

Conover ,W.J. 1980. *Practical Nonparametric Statistics*. John Wiley & Sons. New York.

Cullen, A.C. and H.C. Frey. 1999. Probabilistic Techniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. Plenum Press.

D'Agostino, R. B. and M.A. Stephens. 1986. *Goodness-of-fit techniques*. Marcel Dekker, Inc. New York.

Evans, M., N. Hastings and B. Peacock. 1993. *Statistical Distributions*. Second Ed., John Wiley & Sons. New York.

Filliben, J.J. 1975. The probability plot correlation coefficient test for normality. *Technometrics*. 17(1): 111-117.

Finley, B.L. and D.J. Paustenbach. 1994. The benefits of probabilistic exposure assessment: Three case studies involving contaminated air, water and soil. *Risk Anal*. 14(1): 53-73.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Hostrand Reinhold. New York.

Helsel, D.R. and R.M. Hirsch. 1992. *Statistical Methods in Water Resources*. Elsevier. Amsterdam.

Hoffman, F.O. and J.S. Hammonds. 1992. *An Introductory Guide to Uncertainty Analysis in Environmental and Health Risk Assessment.* ES/ER/TM-35. Martin Marietta.

Hora, S.C. 1992. Acquisition of expert judgment: examples from risk assessment. *J. Energy Eng*. 118(2): 136-148.

Jaynes, E.T. 1957. Information theory and statistical mechanics. *Physical Review*. 106: 620-630.

Johnson, N.L., S. Kotz and N. Balakrishnan. 1995. *Continuous Univariate Distributions*. Volume 2, Second Ed. John Wiley & Sons. New York.

Law, A.M. and W.D. Kelton. 1991. *Simulation Modeling and Analysis.* McGraw-Hill. New York.

Lee, R.C. and Wright, W.E. 1994. Development of human exposure-factor distributions using maximum-entropy inference. *J. Exp. Anal. Environ. Epi*. 4(3): 329-343.

Looney, S.W. and T.R. Gulledge. 1985. Use of the correlation coefficient with normal probability plots. *American Statist*. 39: 297-303.

Mendenhall, W. and R.L. Scheaffer. 1973. *Mathematical Statistics with Applications*. Duxbury Press.

Mood, A.M. and F.A. Graybill. 1963. *Introduction to the Theory of Statistics*. Second Edition. McGraw-Hill, Inc.

Morgan, G.M. and M. Henrion. 1990. *Uncertainty: A guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge University Press. New York.

Nelsen, R.B. 1986. Properties of a one-parameter family of bivariate distributions with specified marginals. *Comm. in Statistics (Theory and Methods)*. 15: 3277-3285.

Nelsen, R.B. 1987. Discrete bivariate distributions with given marginals and correlation. *Comm. in Statistics (Simulation and Computation).* B16: 199-208.

Oregon DEQ. 1998. *Guidance for the Use of Probabilistic Analysis in Human Health Exposure Assessments.* Waste Management and Cleanup Division. Draft. January.

Ott, W.R. 1990. A physical explanation of the lognormality of pollutant concentrations. *J. Air Waste Manag. Assoc*. 40(10): 1378-1383.

Ott, W.R. 1995. *Environmental Statistics and Data Analysis.* CRC Press. Boca Raton.

Roseberry, A.M. and D.E. Burmaster. 1992. Lognormal distributions for water intake by children and adults. *Risk Anal.* 12(1): 99-104.

Royston, J.P. 1982. An extension of Shapiro and Wilk's *W* test for normality to large samples. *Applied Statistics*. 31: 115-124.

Shapiro, S.S. and R.S. Francia, 1972. An approximate analysis of variance test for normality. *J. American Stat. Assoc*. 63: 215-216.

Stanek, E.J. and Calabrese, E.J. 1995. Daily estimates of soil ingestion in children. *Environ. Health Perspect.* 103: 176-285.

Tukey, J.W. 1977. *Exploratory Data Analysis.* Addison-Wesley. Boston.

U.S. EPA. 1982. *Air Quality Criteria for Particulate Matter and Sulfur Oxides*. ECAO, ORD. EPA 600/8-82-029.

U.S. EPA. 1991. *Human Health Evaluation Manual, Supplemental Guidance: "Standard Default Exposure Factors"*. Office of Solid and Emergency Response, Washington, DC. OSWER Directive 9285.6-03.

U.S. EPA. 1992. *Guidance for Data Useability in Risk Assessment, Part* A. Office of Emergency and Remedial Response, Washington, DC. Publication 9285.7-09A.

U.S. EPA. 1994. *Guidance for Conducting External Peer Review of Environmental Regulatory Models.* Office of the Administrator, Washington, DC. EPA/100/B-94-001. July.

U.S. EPA. 1996. *Summary Report for the Workshop on Monte Carlo Analysis.* Risk Assessment Forum. EPA/630/R-96/010.

U.S. EPA. 1997a. *Exposure Factors Handbook.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fa, Fb, and Fc.

U.S. EPA. 1997b. *Use of Probabilistic Techniques (Including Monte Carlo Analysis) in Risk Assessment*, Memorandum from Deputy Administrator Hansen and *Guiding Principles for Monte Carlo Analysis.* EPA/630/R-97-001.

U.S. EPA. 1999a. *Report of the Workshop on Selecting Input Distributions for Probabilistic Assessments.* Risk Assessment Forum. EPA/630/R-98/004. January.

U.S. EPA. 1999b. *Options for Development of Parametric Probability Distributions for Exposure Factors.* Office of Research and Development. Research Triangle Institute Final Report. April 6.

U.S. EPA. 1999c. *Development and Evaluation of Probability Density Functions for a Set of Human Exposure Factors.* Office of Emergency and Remedial Response. University of California Draft Report. April.

Vose, D. 1996. *Quantitative Risk Analysis: A Guide to Monte Carlo Modelling*. John Wiley & Sons. New York.

CHAPTER 4 (PART 1 OF 2)

USING PROBABILISTIC ANALYSIS IN HUMAN HEALTH RISK ASSESSMENT

4.0 INTRODUCTION

This chapter outlines the basic concepts associated with the use of probabilistic analysis in human health risk assessments. At this time, national policy recommends that probabilistic analysis for human health assessments generally should be confined to the exposure variables (i.e., concentration, exposure frequency, exposure duration, etc.). Therefore, probabilistic analysis for human health risk assessments generally should not model variability or uncertainty in the cancer slope factor (CSF), reference dose (RfD), or reference concentration (RfC).

This chapter provides guidance on characterizing variability and uncertainty in exposure variables. A discussion about variability and uncertainty in the concentration term is also included in Section 4.1. To aid in the development and review of probabilistic risk assessments, examples of input assumptions are provided. A recommended checklist for reviewing a PRA is provided in Section 6.2, Table 6-1.

4.1 CHARACTERIZING VARIABILITY IN EXPOSURE VARIABLES

The general equation (Exhibit 4-1) used for calculating exposure (i.e., average daily intake) is shown in the accompanying text box. In PRA, the only modification is that probability distributions rather than single values (i.e., point estimates) are specified for one or more variables to characterize variability or uncertainty (see Chapter 3). A Monte Carlo simulation is executed by repeatedly selecting random values from each of these distributions and calculating the corresponding exposure. Each input can be verified by closely examining the results of the simulation. If the complete set of random values for an exposure variable are sorted and plotted on a graph, the resulting frequency distribution should provide a close approximation of the original distribution that was specified.

EXHIBIT 4-1

GENERAL EQUATION FOR ESTIMATING EXPOSURE TO A SITE CONTAMINANT

$$I = \frac{C \cdot CR \cdot EF \cdot ED}{BW \cdot AT} \quad \text{Eq. 4-1}$$

where,

I	=	daily intake
C	=	contaminant concentration
CR	=	contact rate (ingestion, inhalation, dermal contact)
EF	=	exposure frequency
ED	=	exposure duration
BW	=	body weight
AT	=	averaging time

EXHIBIT 4-2

DEFINITIONS FOR CHAPTER 4

95% UCL for mean - The 95 percent upper confidence limit (UCL) for a mean of a population is defined as a value that, when repeatedly calculated for randomly drawn subsets of size (n), equals or exceeds the true population mean 95% of the time. Although the 95% UCL provides a conservative estimate of the mean, it should not be confused with a 95th percentile. As sample size increases, the difference between the UCL for the mean and the true mean decreases, while the 95th percentile of the distribution remains relatively unchanged, at the upper end of the distribution.

95th percentile - The number in a distribution such that 95% of the values in the distribution are less than the number and 5% are greater.

ARARs - Applicable or relevant and appropriate requirements. The NCP (U.S. EPA, 1990) states that ARARs shall be considered in determining remediation goals. A maximum contaminant level (MCL) from the Safe Drinking Water Act is an example of an ARAR. If an ARAR meets the requirements of the NCP for protectiveness, it may be selected as a site-specific cleanup level.

Arithmetic Mean (AM) - A number from a distribution or sample that is the average of all members of the sample. Usually obtained by summing all the members in the sample and dividing by the number of members.

Assessment Endpoint - The specific expression of the population or ecosystem that is to be protected. It can be characterized both qualitatively and quantitatively in the risk assessment.

Bayesian Analysis - Also called subjective probability. These methods start with a description of the probability of an event as the degree of belief or confidence in that event occurring, given some state of knowledge. Bayesian methods combine subjective probabilities with data to arrive at posterior probabilities. Bayesian analysis provides a way of integrating professional judgement with data in a rigorous mathematical framework (Appendix E).

Box Plot - Graphical representation showing the center and spread of a distribution, sometimes with a display of outliers. This guidance uses boxplots to represent the following percentiles, where the box represents the 25th and 75th percentiles, and the hinge represents the 5th and 95th percentiles. Other definitions for the box and hinge may be used.

Central Tendency Risk/ Central Tendency Exposure (CTE) - A risk descriptor representing the average or typical individual in the population, usually considered to be the arithmetic mean or median of the risk distribution.

1

EXHIBIT 4-2 (CONT'D)

DEFINITIONS FOR CHAPTER 4

Parameter - value that characterizes a feature of a population. Parameters that define a probability distribution for a random variable commonly characterize the location, scale, shape, or bounds of the distribution. For example, a truncated normal probability distribution may be defined by four parameters: arithmetic mean [location], standard deviation [scale], and min and max [bounds]. It is important to distinguish between an exposure variable (e.g., ingestion rate) and a parameter (e.g., arithmetic mean ingestion rate).

Reasonable Maximum Exposure (RME) - The highest exposure that is reasonably expected to occur at a site (U.S. EPA, 1989a). The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.

Figure 4-1 shows a typical example of an input distribution for drinking water ingestion rate. In the graph, the height of the bars represents the relative frequency of ingestion rates in the population and the spread of the bars is the varying amounts of water ingested per day. The frequency distribution was generated by randomly sampling (10,000 times) from a lognormal probability distribution function (PDF); the type or shape of the distribution is set by the user. When the frequency distribution is overlain with the lognormal PDF, it becomes clear that 10,000 iterations yields a close approximation of the theoretical distribution. The difference between the frequency distribution and the parametric distribution will approach zero as the number of random samples approaches infinity. With the speed of modern computers, 10,000 random samples can be selected in a minute. Guidance on selecting and fitting distributions to data is presented in Chapter 3.

Each probability distribution used in a Monte Carlo analysis gen should be presented with sufficient detail that the analysis can be reproduced. This information may be presented in tabular and/or graphical summaries. Important information for a summary table would include a description of the distribution type (e.g., lognormal, gamma, etc.), the parameters that define the distribution, a discussion on how the distribution was selected and the representativeness of the underlying data relative to the exposures being evaluated. If other alternatives were available, the report should discuss why a particular selection was made. It should also be made clear if the distribution describes variability or uncertainty. Graphical summaries of the distributions may include both PDFs and CDFs, and should generally be used to document distributions that characterize site-specific data. The decision to exclude graphical summaries for some of the more common distributions should be made in consultation with the Agency risk assessor.

The calculation of risk involves sampling from each of the exposure variable distributions and then incorporating toxicity; this process results in an output of risk that is also a distribution. When the

calculation of risk (or any other model endpoint) is repeated many times using Monte Carlo techniques to sample the variables at random, the resulting distribution of risk estimates can be displayed in a similar fashion. The type of summary graph used to convey the results of a Monte Carlo analysis depends on the risk management needs. For example, Chapter 1 (Figure 1-2) illustrates the histogram of risk estimated from a Monte Carlo simulation, and the corresponding PDF for risk. The height of each bar indicates the relative probability of a given risk value. This type of summary can effectively illustrate the relationship between the point estimate of RME risk, which can be placed on this graph, and the probabilistic estimate of the high-end risk range.

In addition, the Cumulative Distribution Function (CDF) can be especially informative for illustrating the percentile corresponding to a particular risk level of concern (e.g., $1x10^{-6}$). Figure 4-2 illustrates both the PDF and CDF for risk. Factors to consider in deciding whether to display the PDF or CDF are discussed in Chapter 1. When in doubt about the appropriate type of summary to use, provide both the PDF and CDF for all risk distributions. At a minimum, each summary output for risk should provide the risk descriptors of concern (e.g., 50^{th}, 90^{th}, 95^{th}, and 99.9^{th} percentiles of the PRA; CTE and RME of the point estimate risk assessment).

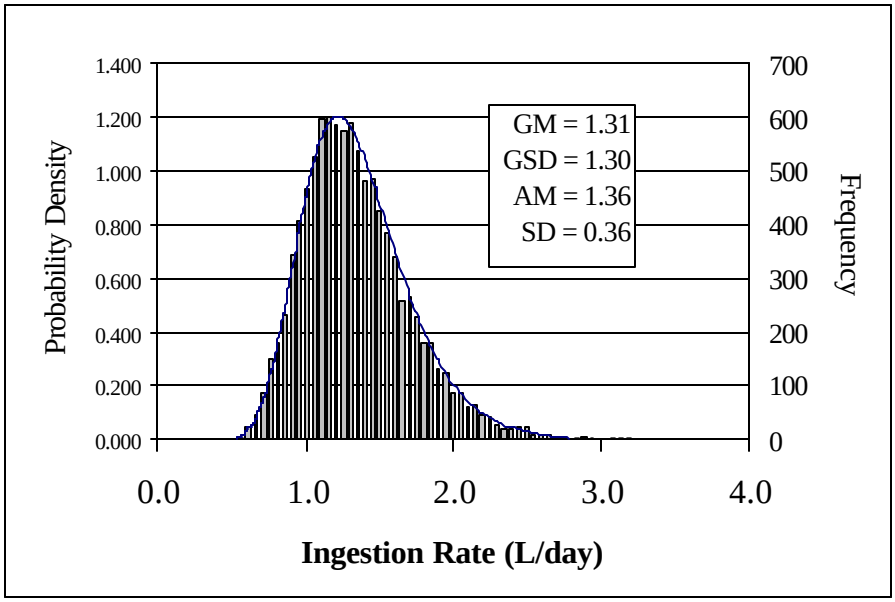


Figure 4-1. Results of Monte Carlo simulation yielding a frequency distribution from 10,000 random samples based on a theoretical lognormal PDF (parameters given on graph). The distribution represents inter-individual variability in adult drinking water intakes and is characterized by two parameters - typically Log ~(GM, GSD) or Log ~(AM, SD).

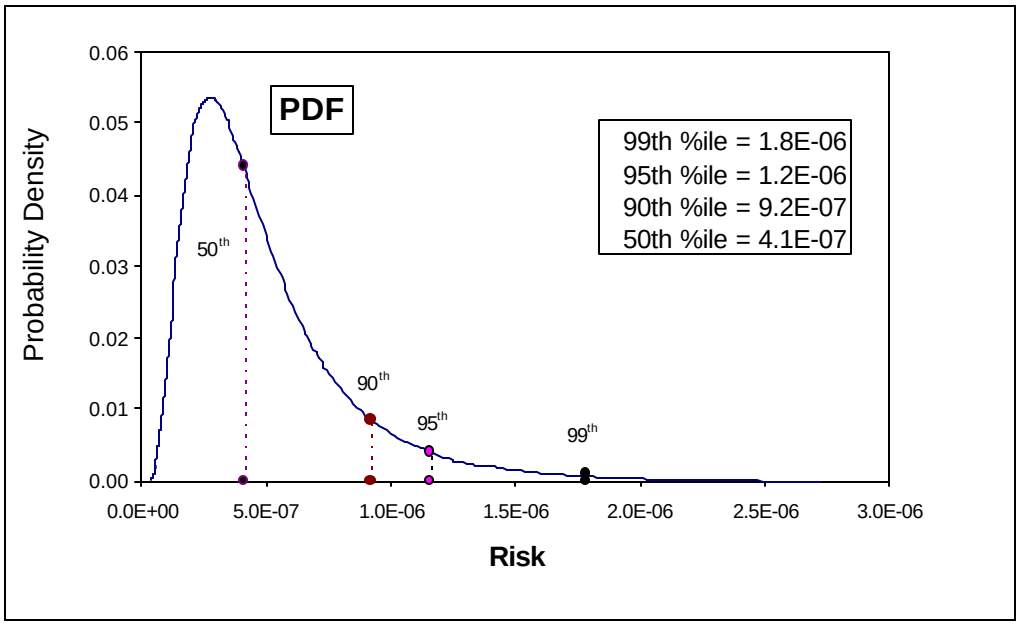


Figure 4-2 (part 1 of 2). Hypothetical PRA results showing a probability density function (PDF, top panel) and cumulative distribution function (CDF, bottom panel) for cancer risk with selected summary statistics for central tendency and high-end percentiles. The PDF rises to a maximum of 0.055 at a risk of about 5E-07. The CDF rises to a maximum cumulative probability of 1.0. The CDF clearly shows that the level of regulatory concern of $1x10^{-6}$ falls between the 90th and 95th percentiles of the risk distribution.

CHAPTER 4 (PART 2 OF 2)

USING PROBABILISTIC ANALYSIS IN HUMAN HEALTH RISK ASSESSMENT CONTINUED

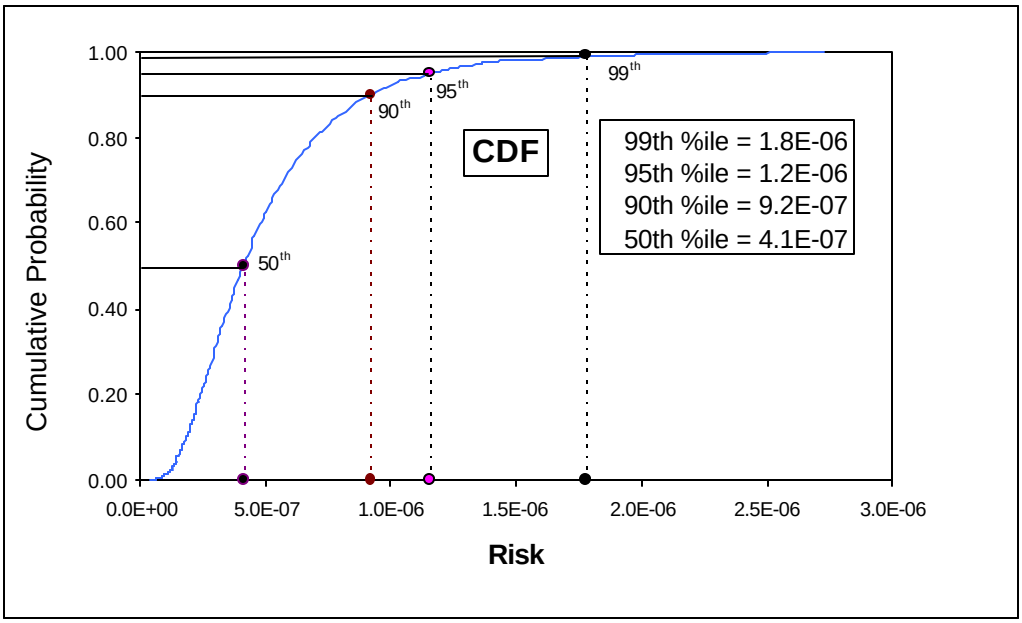


Figure 4-2 (part 2 of 2). Hypothetical PRA results showing a probability density function (PDF, top panel) and cumulative distribution function (CDF, bottom panel) for cancer risk with selected summary statistics for central tendency and high-end percentiles. The PDF rises to a maximum of 0.055 at a risk of about 5E-07. The CDF rises to a maximum cumulative probability of 1.0. The CDF clearly shows that the level of regulatory concern of $1x10^{-6}$ falls between the 90th and 95th percentiles of the risk distribution.

4.2 CHARACTERIZING UNCERTAINTY IN EXPOSURE VARIABLES

PRA can also provide quantitative information on the confidence associated with a risk estimate to assist the risk manager in choosing an appropriate percentile in the RME range of the estimated risk distribution. This choice will consider the uncertainty in the risk estimate (see Chapter 7). Most site-specific PRAs will focus on the variability in exposure among a population to arrive at a distribution of risks for that population. From the distribution of risks for the population, the risk to a single hypothetical individual generally will considered in making a remedial decision.

Questions will almost certainly arise concerning the confidence or certainty around risk estimates. For example, how much confidence do we have that enough soil samples were collected to adequately characterize contaminant concentrations on site? In general, probability distributions are used to characterize either variability or uncertainty in an exposure variable. The PDF should capture the nature of the exposure variable. Some aspects to consider in selecting and fitting PDFs are presented in Chapter 3.

A more complex Monte Carlo simulation technique, which distinguishes variability from uncertainty in risk estimates, is two dimensional Monte Carlo analysis (2-D MCA) (see Appendix E). Figure 4-3 provides hypothetical results of a 2-D MCA where a confidence interval has been quantified around a 95^{th} percentile of variability in risk (also see Chapter 2, Figure 2-2). In Areas 1 and 3, the confidence intervals are fairly narrow, which suggests a high degree of confidence that the risks in Area 1 are negligible, and the risks in Area 3 may require some type of action. The fairly broad confidence intervals around the risk estimates for Areas 2 and 4 give us less confidence in the results. Of course, the level of confidence in the risk estimate should be considered and may affect the risk manager's decision.

2D-MCA cannot be applied at all sites. However, this simulation technique may be useful when a risk assessor has sufficient data and/or a clear understanding of one or more exposure variables such that uncertainty may be quantified. As discussed in Chapter 3, often there is uncertainty in the probability distribution that best describes the variability in an exposure factor. In such cases, this guidance strongly recommends not using 2-D MCA to select from the candidate distributions at random (i.e., according to ad hoc probabilities). Rather, separate 1-D MCA simulations should be run with the different candidate models, and the risk distribution from each should be compared to evaluate the importance of this source of uncertainty. In general, 2-D MCA should be used to propagate parameter uncertainty in probability distributions for multiple exposure factors rather than explore uncertainty in the choice of probability distributions used to characterize variability.

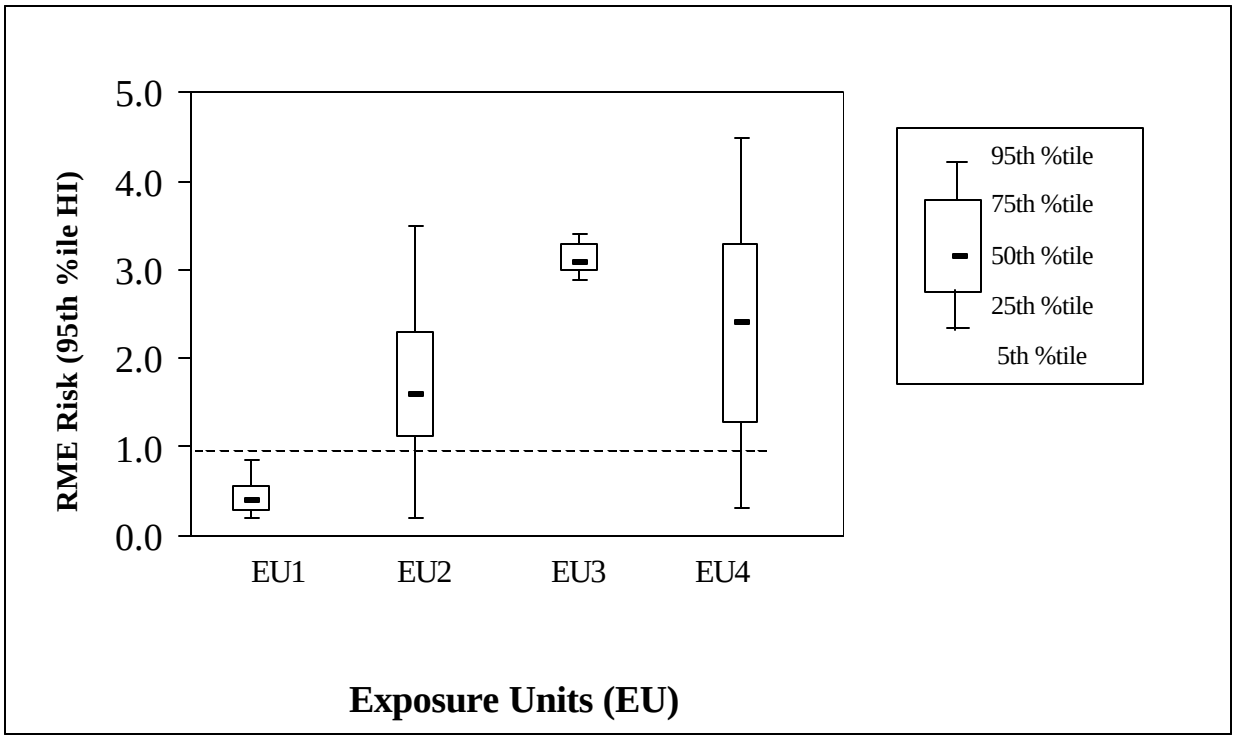


Figure 4-3. Box and whisker plots characterizing uncertainty in the RME risk estimates (95th percentile of the Hazard Index) at four locations. The box represents the interquartile range (25th to 75th percentiles) while the whiskers represent the 90% confidence limits (5th to 95th percentiles).

4.3 CHARACTERIZING VARIABILITY AND UNCERTAINTY IN THE CONCENTRATION TERM

Often, an important source of variability and uncertainty in risk assessment is the concentration of one or more contaminants in the various exposure media. When specifying either a point estimate or a distribution for the concentration term, it is important to distinguish between variability and uncertainty. In any risk assessment, the concentration term will necessarily reflect an assumption about (1) the spatial and temporal variability in contamination and (2) the behavior of the receptor.

The concentration term is linked to the concept of an exposure unit (i.e., the geographical area in which a receptor exists and may be exposed to a contaminated medium during the time period of interest). Environmental sampling will provide information about the contamination within an exposure unit. Site-specific information regarding activities of receptors may guide assumptions about the receptor's contact with exposure media.

- *An exposure unit in a Superfund risk assessment is the geographical area about which a receptor moves and contacts a contaminated medium during the period of the exposure duration. Either random or non-random movement and contact can be modeled.*

The size of the exposure unit should be appropriate for the receptor being considered. For example, depending on the receptor's movement and activities, an exposure unit may be as small as a child's play area (e.g., sand box) or as large as an ecological receptor's territory size (e.g., industrial facility). In some cases, it may be appropriate to define multiple exposure units based on the receptor's activity patterns, the chemicals of potential concern, and the potential exposure media. For example, residential exposures to children may reflect exposures via soil and dust ingestion both at the primary residence and at a day care facility.

Over a short period of time (e.g., days), contaminant concentrations contacted by a receptor are likely to vary depending on the spatial variability of contamination and the movements of the receptor. Similarly, different individuals may be exposed to different concentrations based on inter-individual variability in activity patterns. If information regarding activity patterns is unavailable, receptors are typically assumed to exhibit random movement such that there is an equal probability of contacting any area within the exposure unit. In addition, in Superfund risk assessments, the toxicity criteria are often based on health effects associated with chronic exposure (e.g., lifetime risk of cancer following chronic daily intake over a period of 30 years). Hence, the most appropriate expression or modeling approach for the concentration term is one that characterizes the long-term average exposure point concentration within the exposure unit.

The concepts of long-term exposure and random access are most easily envisioned when soil is the contaminated medium. One can think of the receptor moving within the exposure unit while the soil remains fixed. With groundwater, exposure occurs at a fixed point in space (the wellhead) and concentrations may change with time. With consumption of contaminated fish, both the contaminated medium and the exposure point change throughout the exposure duration. For example, fish populations can change over time and anglers may fish from different locations within a lake. Nonetheless, the long-

term random exposure, as represented by an arithmetic mean, may be applied to any environmental medium to derive a concentration term. Long-term sampling, fate-and-transport modeling or other techniques may be used to evaluate present and future contaminant concentrations.

Characterizations of the concentration term in both point estimate and probabilistic risk assessment generally focus on uncertainty. The numerous potential sources of uncertainty in the estimate of the true mean concentration can be grouped into the following two broad categories:

1. Uncertainties in the sample data. A limited number of measurements in the sample are used to make inferences about the true arithmetic mean concentration and the spatial distribution of concentrations at a site. Uncertainties may arise from many factors, including both sampling variability and measurement error.

2. Uncertainties in the behavior of the receptor. Even in the case of extremely well characterized sites, it remains uncertain whether the receptor will contact the environmental medium in a temporal and/or spatial distribution that can be adequately represented by the environmental samples collected.

The correspondence between the concentration contacted by the receptor and that represented by sampling is a function of both the site characterization and the knowledge of the behavior of the receptor. A lack of knowledge in both categories accounts for the uncertainty in the concentration term.

4.3.1 SPECIFYING A DISTRIBUTION FOR THE CONCENTRATION TERM WHEN EXPOSURE IS RANDOM

If a receptor's contact with the contaminated medium within an exposure unit is truly random, then the receptor can be expected to contact all parts of the contaminated medium for an equal amount of time. In this case, the long-term average exposure concentration experienced by the receptor will be adequately represented by the arithmetic mean. This is true regardless of the underlying distribution of the environmental sampling data (e.g., lognormal, normal, beta, mixed). Inter-individual variability in the long-term average exposure concentration within the receptor population will be minimal under these assumptions.

Methods for characterizing uncertainty are summarized in Table 4-1. The methods are briefly discussed in this chapter and presented in more detail in Appendix D.

Single Point Estimates for Uncertainty

In conventional point estimate risk assessments, uncertainty in the concentration term is characterized by the 95% UCL for the arithmetic mean. EPA's *Supplemental Guidance to RAGS: Calculating the Concentration Term* (U.S. EPA, 1992) recommends using the Student-t statistic to calculate the UCL for a normal distribution, and the H-statistic when the data set appears lognormally distributed. Alternatively, statistical resampling techniques, such as the bootstrap and jacknife, can be used without selecting and fitting a probability distribution to the sample (U.S. EPA, 1997a). These techniques are sometimes referred to as "distribution free", and add considerable flexibility by characterizing uncertainty

in any summary statistic for the concentration term (e.g., arithmetic mean, median, or selected percentiles). Parametric and nonparametric bootstrap techniques are generally not a substitute for good site sampling (refer to data quality objectives guidance, U.S. EPA, 1993b). A small sample will high variance will yield poor parameter estimates, especially for high-end percentiles (e.g., > 90^{th} percentile), no matter which technique is used. Each of these methods of calculating the 95% UCL apply equally to PRA, and may provide measures of parameter uncertainty in the risk model.

Table 4-1. Methods for characterizing uncertainty in the concentration term.

Approach	Example of Model Input	Primary Use in PRA	Example of Model Output
Single Point Estimate	• 95% UCL	1-D MCA[1]	• PDF for variability in risk, calculated using the 95% UCL for the concentration term
Multiple Point Estimates	• 95% LCL • sample mean • 95% UCL	1-D MCA[1]	• Three PDFs for variability in risk, representing the 90% confidence interval for each percentile of the risk distribution. The 90% CI *only accounts for uncertainty in the concentration term*, not any other exposure variables.
Parametric Probability Distribution for Uncertainty	Triangular PDF with • min = 95% LCL • mode = sample mean • max = 95% UCL	2-D MCA	• A family of PDFs for variability in risk from which the 90% confidence interval can be calculated for each percentile of the risk distribution. The 90% CI reflects uncertainty in *one or more variables*, including the concentration term
Nonparametric Probability Distribution for Uncertainty	Bootstrap estimate of the uncertainty distribution (see Appendix D)	2-D MCA	• same as parametric probability distribution for uncertainty

[1] The probability distribution(s) for variables (other than the concentration term) may characterize either parameter uncertainty (for a 1-D MCA of uncertainty) or variability (for a 1-D MCA of variability).

Multiple Point Estimates for Uncertainty

In general, when characterizing variability and uncertainty simultaneously in PRA, RAGS Vol. 3 recommends using a 2-D MCA (see Section 4.2). This approach allows probability distributions for variability and probability distributions for parameter uncertainty to be characterized and propagated separately in a Monte Carlo model. The box plots for HQ shown in Figure 4-3 are examples of the output from a 2-D MCA. While this type of information can be useful for making risk management decisions, as discussed in Chapter 1 (Figure 1-4), 2-D MCA may involve additional time and effort that may not be warranted for every site.

If parameter uncertainty in a single variable is of interest, multiple 1-D MCA simulations can yield the same results a 2-D MCA simulation. The advantage of running multiple 1-D MCA simulations is that it usually takes less time and effort. If parameter uncertainty for more than one variable is of interest, multiple point estimate calculations would not be sufficient to model the numerous possible combinations of input values; 2-D MCA would be the recommended choice. In practice, while there is uncertainty with each parameter estimate, sensitivity analyses (Chapter 2) may reveal that most of the uncertainty is attributable to a single parameter, such as the concentration term.

Using a hypothetical example, assume that a risk assessor wishes to estimate the 90% confidence interval for the 95th percentile risk based on uncertainty in the true mean concentration. For a 2-D MCA, the 90% confidence interval corresponds with the interval on the distribution for uncertainty in the risk estimate that is bounded by the 5th and 95th percentiles. Since there is only one source of parameter uncertainty, the 5th and 95th percentiles for uncertainty in risk will directly correspond with the 5th and 95th percentiles for uncertainty in the concentration term. The risk distribution obtained with the mean will represent the most likely risk estimate and the distributions obtained with the upper and lower bounding values will represent the 90% upper and lower confidence limits of the risk distribution. Percentiles of the distribution for uncertainty in the mean concentration can be estimated from bootstrap techniques, or from the 95% LCL and UCL for the mean concentration as described above.

Parametric Probability Distribution for Uncertainty

If multiple sources of uncertainty are propagated in a Monte Carlo analysis (e.g., 1-D MCA for uncertainty; or 2-D MCA for variability and uncertainty), then it may be useful to characterize uncertainty in the mean concentration with a probability distribution rather than multiple point estimates. For samples that are fit to a normal or lognormal distribution, estimates of the 95% LCL and 95% UCL may provide a reasonable measure of the range of uncertainty in the concentration term. Given this range, and the arithmetic mean concentration, a risk assessor may wish to use a triangular distribution to characterize uncertainty, as shown in Figure 4-4. This choice of distributions provides a reasonable screening tool for sensitivity analysis. In addition, based on Maximum Entropy principles (Chapter 3, Section 3.4.6), the triangular distribution would be the preferred PDF for maximizing uncertainty given estimates of the plausible range and most likely value for the arithmetic mean. If parameter uncertainty, such as uncertainty in the arithmetic mean concentration, is determined to be a major source of uncertainty in the risk estimates, additional methods may be explored, including extending the bounds of the uncertainty

distribution (e.g., 97% Confidence Interval, given by the 97.5% LCL and UCL), and collecting or researching additional data to reduce uncertainty (i.e., Tier 2 of Figure 1-3).

Probability distributions for variability and uncertainty should not be combined in a 1-D MCA simulation. For example, assume a triangular distribution for uncertainty in mean concentration was combined with distributions representing variability in other exposure variables using 1-D MCA. The result would be a single distribution for risk that reflects both uncertainty and variability. In such a case, distinguishing the variability in risk from the uncertainty in the risk estimates would not be possible. EPA's *Guiding Principles for Monte Carlo Analysis* recommends against mixing distributions of variability and uncertainty in a 1-D MCA (U.S. EPA, 1997b) to avoid such ambiguities.

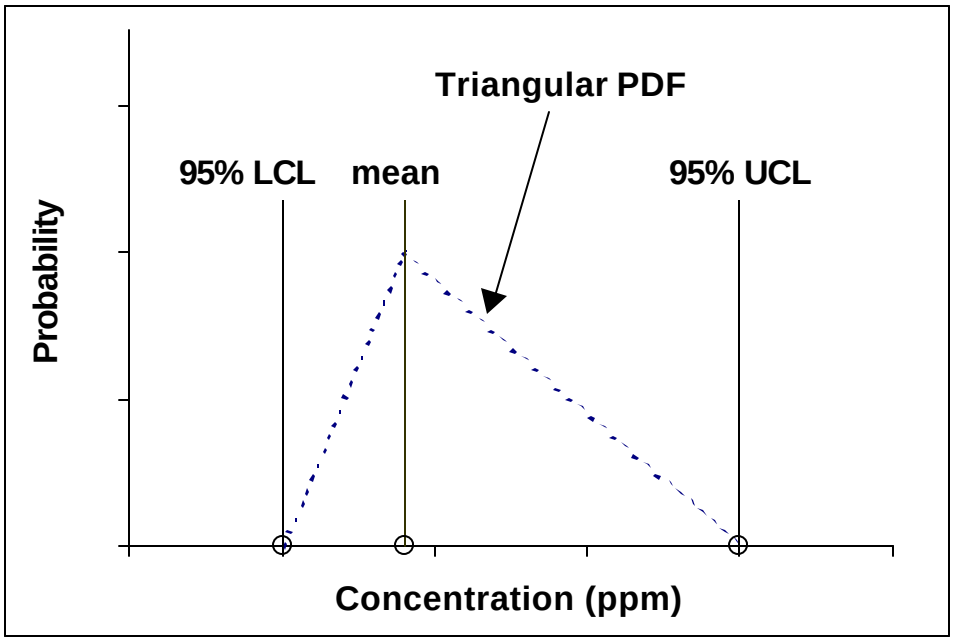


Figure 4-4. Uncertainty in the concentration term showing hypothetical estimates of the 95% LCL, sample mean, and 95% UCL. For the point estimate method, three 1-D MCA simulations could be run, each simulation using one of the 3 point estimates representing uncertainty. For a 2-D MCA, a triangular distribution for uncertainty may be fit to the three points. Note that in this example, the mode (most likely value) of the triangular distribution is defined by the sample mean in order to reflect the assumption that this is the best estimate for the long-term exposure point concentration.

Nonparametric Probability Distribution for Uncertainty

Parametric and nonparametric bootstrap techniques are relatively straight forward to implement, and can be used to characterize the entire probability distribution for uncertainty in the mean concentration (see Appendix D). As with the parametric probability distribution approach described above, bootstrapped estimates of the mean are useful when multiple sources of uncertainty are propagated in a Monte Carlo model (e.g., uncertainty in concentration, exposure duration, etc.).

4.3.2 SPECIFYING A DISTRIBUTION FOR THE CONCENTRATION TERM WHEN EXPOSURE IS NOT RANDOM

For some situations, the assumption of random exposures may be clearly incorrect. A quantitative estimate of variability in the concentration term may be informative. For example, groundwater concentration measurements may show a large variance when sampled from wells in different locations (Figure 4-5). Typically, residential receptors do not sample randomly from different wells, but draw chronically from individual wells; the contacted wells in an area may have contaminant concentrations that vary with location of the well. In such a case, the exposure unit is a single wellhead.

A probabilistic analysis of variability may be used to account for temporal variability due to seasonal fluctuations in the water table. Other factors may also influence the fate and transport of the chemicals and, hence, the concentration. The risk assessor should consult extensively with a hydrogeologist to obtain an appropriate estimate of the long-term average concentration at a given location.

Temporal variability in the concentration term may be an important component of individual variability in exposures, depending on the chemical and the exposure medium. For example, wind erosion may change chemical concentrations in surface soil. Leaching may change concentrations in both subsurface soil and groundwater. Large predatory fish may have high concentrations of contaminants due to bioaccumulation, and may be preferentially selected by an angler for consumption. Such factors generally should be considered early in the risk assessment process and included in the conceptual site model.

Simulations of long-term average intake based on a series of short-term doses may be useful; for example, in simulations of groundwater concentrations which fluctuate over time. The Microexposure Event (MEE) approach simulates long-term intake as the sum of individual exposure events and is appropriate for this situation (Appendix E). The time step for MEE is an important consideration and will depend on the rate of change of the most rapidly changing exposure variable. In addition, there should be a correspondence between the time periods over which data were obtained and the time step used in the MEE model.

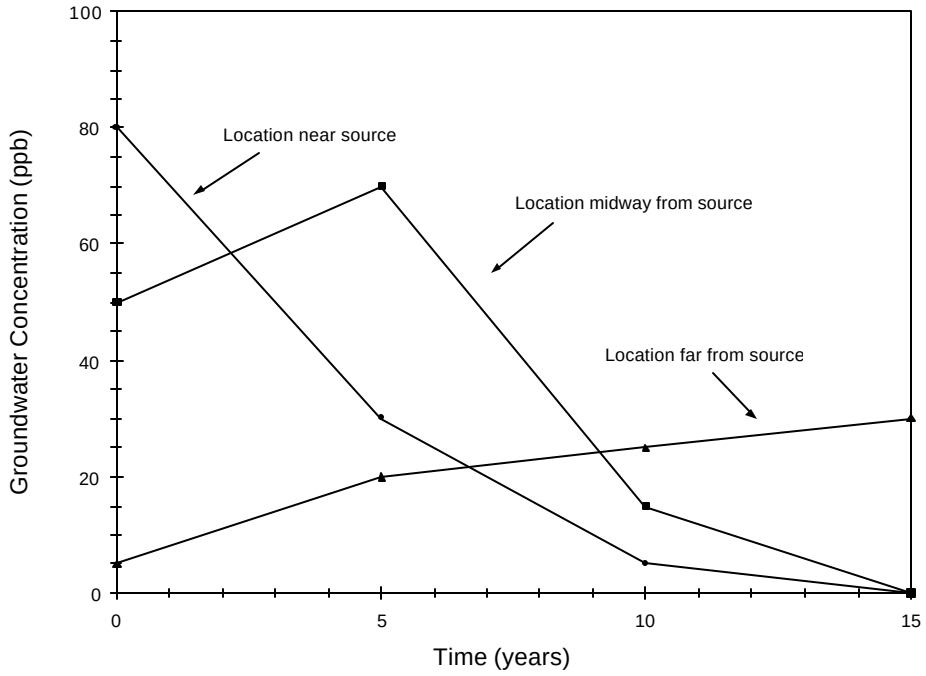


Figure 4-5. Spatial and temporal variability in contaminant concentrations in groundwater. Fluctuations in the groundwater plume will depend on the hydrogeology of the site as well as the seasonal fluctuations in the water table. In this hypothetical example, concentrations are declining over time at distances nearest to the source, and concentrations are increasing as the plume moves farther from the source.

4.3.3 SPECIFYING A DISTRIBUTION FOR THE CONCENTRATION TERM WHEN EXPOSURE IS NOT LONG TERM

At times, acute or short-term exposures may be distinguished from chronic exposure scenarios and warrant a separate assessment. Acute exposures can be defined differently for different sites. It may refer to intermittent exposure events over long periods (inhalation exposures from temporary incinerator emissions; landscaping), or multiple exposure events over a short period (e.g., construction workers renovating a building). The frequency and duration of the exposure events will be different than that of chronic exposure scenarios. In addition, the exposure unit may represent the specific areas of the site that the receptor is expected to contact during the short period. Contact within this exposure unit is still assumed to be random, unless information on time-use and activity patterns suggest that exposures are likely to be greater within certain locations of the exposure unit (e.g., hotspots; confined areas). The parameter of interest is still the mean exposure concentration (or 95% UCL). The use of a distribution defining lower and upper confidence limits on the mean concentration within the exposure unit to represent uncertainty in the concentration term, as discussed in Section 4.3.1, would still be appropriate. A possible exception might be for acute exposure scenarios in which contaminant levels from a limited number of sample points (or even a single sample point) may pose an immediate health risk. In these instances, the use of a point estimate for the concentration term may be more appropriate.

4.4 EVALUATING THE RISK RANGE AND THE RME RANGE

The results of a PRA will generally yield a probability distribution for risk (or Hazard Index), from which the risk corresponding to the CTE and RME can be identified. In general, a risk manager will identify an RME from the high end of the distribution (i.e., the 90^{th} to 99.9^{th} percentiles) for risk (or Hazard Index). As discussed in Chapter 1 (see Section 1.4), for clarity in this guidance, the 90^{th} to 99.9^{th} percentiles of the risk distribution are collectively referred to as the recommended *RME range*. Therefore, in order to utilize PRA results to establish that a cleanup goal is sufficiently protective, two questions will generally need to be addressed:

1) How will the RME risk be identified from the RME range of the risk distribution? and
2) How will information on uncertainty in the high-end risk estimates be utilized in this process?

The NCP (U.S. EPA, 1990) discusses a generally acceptable range for cumulative excess cancer risk of 10^{-6} to 10^{-4} for protecting human health. Furthermore, the NCP specifies 10^{-6} as a point of departure for determining remediation goals when ARARs are not available or sufficiently protective. The range, 10^{-6} to 10^{-4}, will be referred to as the *risk range* in this guidance. If the Hazard Index (HI) is greater than 1, there may be a concern for potential non-cancer effects. Generally, where cumulative carcinogenic risk to an RME individual is less than 10^{-4}, and the non-carcinogenic hazard quotient is less than or equal to 1, action is not warranted unless there are adverse environmental impacts or ARARs (e.g., maximum contaminant levels) are exceeded (U.S. EPA, 1991).

In point estimate risk assessment, the RME is estimated from the combination of point estimates appropriate for the RME. In PRA, a risk manager chooses a percentile from the RME range to represent

the RME. This section provides guidance on factors to base the choice of RME from within the RME range. Additional discussion is provided in Chapter 5.

4.4.1 RISK ASSESSMENT CONSIDERATIONS

Much like the data quality objectives that are evaluated during sample design and collection, it is important to consider the quality of the data that are used in the risk assessment (i.e., qualitative and quantitative uncertainties) in order to evaluate the strengths and weaknesses of the assessment (U.S. EPA, 1993b). Communication between risk managers, risk assessors, and other technical team members is vital at this stage. The main question to be answered is, "How well do the inputs to the risk assessment represent exposure pathways and behaviors at a given site?". The answer to this question can be expressed qualitatively (e.g., high, medium, or low) or quantitatively (e.g., 2-D MCA). For example, are the inputs based on the collection of site-specific data or are they based on generic national data? If based on generic data, how closely do they represent the actual exposure factors at a given site? Also, how certain are the toxicity benchmarks (i.e., do they have high, medium, or low confidence, reflecting low, medium or high uncertainty)? The importance of data representativeness in selecting and fitting distributions is discussed in Chapter 3 (Section 3.4).

A specific example of potentially poor representativeness would be the use of national data for estimating the exposure frequency of adult workers when the receptor of concern is a railroad worker. Such workers may typically be on the site for only 100 days/year. If the risk assessment were based on the national default assumption of 250 days/year, this choice would give a high bias to the risk estimate. When evaluating the representative of a distribution, care should be given to characterizing both current and future exposure scenarios.

Another example of a site-specific exposure factor that may vary considerably among different locations is fish ingestion rates. At sites where ingestion of fish contaminated with metals poses a concern, tissue concentrations from fish fillets collected on site are often used to determine the concentration term. However, a cultural practice of people harvesting fish on site may include consuming some of the internal organs of the fish in addition to the fillets. If the metal contaminants selectively accumulate in the internal organs instead of the fillet tissues, use of data on fillets and omitting data on internal organ contaminants would give a low bias to the risk estimate.

This type of evaluation of the data can provide the risk assessor and risk manager a qualitative indication of whether the data used in the risk assessment are likely to over- or underestimate the risk. Such qualitative evaluation is invaluable. If there is substantial qualitative uncertainty in the risk assessment, a prudent course of action for risk managers would generally be to select a percentile toward the upper end of the RME range (i.e., 90^{th} to 99.9^{th} percentiles) or to collect additional data when the direction of the bias is unknown.

PRA can also provide quantitative information on the confidence associated with a risk estimate, which could assist the risk manager in choosing an appropriate percentile in the RME range. Most site-specific PRAs will focus on the variability in exposure parameters in a population to arrive at a distribution

of risks for that population or any given individual in that population. However, questions may arise concerning the confidence or certainty around those risk estimates. For example, how much confidence do we have that enough soil samples were collected to adequately characterize contaminant concentrations on site? A more complex PRA method that separates variability from uncertainty, such as a 2-D MCA, could be used to quantify the uncertainty or confidence surrounding the contaminant concentration and, subsequently, the final risk estimate (see Figure 4-3 and Appendix E).

4.4.2 SITE-SPECIFIC FACTORS

The demographic and behavioral features of a potentially exposed population and the physical and geographical factors at a site can increase or decrease exposure to contaminated media. Examples of such factors are listed in Exhibits 4-3 and 4-4. These factors should be considered in defining exposure pathways and characterizing exposure variables in the risk assessment. Such site-specific information may support a decision to evaluate the entire RME range (90^{th} to 99.9^{th} percentile) before selecting the percentile that represents RME risk. A departure from the 95^{th} percentile would depend on whether or not qualitative factors, or factors not quantitatively addressed in the exposure assessment, suggest an increased or decreased exposure, and hence, risk. In practice, multiple and sometimes competing factors will need to be balanced in order to determine an appropriate percentile for the RME risk (see hypothetical example in Section 4.4.7). When evaluating such factors, risk assessors should account for both current and future land use scenarios.

EXHIBIT 4-3

EXAMPLES OF DEMOGRAPHIC, CULTURAL, AND BEHAVIORAL FACTORS THAT CAN AFFECT EXPOSURE

- subsistence fishing, hunting, or ingestion of home-grown produce
- exposures to cultural foods or medicines that contain contaminants
- preparation of foods in containers that contain contaminants
- hobbies and other personal practices resulting in exposure to contaminants
- age of the population (e.g., children may have greater exposure and susceptibility than adults; U.S. EPA, 1995a; 1996)

Subpopulations may be at increased risk from chemical exposures due to increased sensitivity, behavior patterns that result in high exposures, and/or current or past exposures from other sources. Environmental health threats to children are a particular concern (U.S. EPA, 1995a, 1996). Once identified, a subgroup can be treated as a population in itself, and characterized in the same way as the larger population using similar descriptors for population and individual risk (U.S. EPA, 1995b).

4.4.3 BIOLOGICAL DATA

Biological monitoring data and/or other biomarker data can be useful sources of information for evaluating uncertainty in an exposure assessment. In human health risk assessments, these data can provide an indication of the magnitude of current or past exposures and the degree to which the exposures are correlated with contaminated media. Examples of biological data are lead in blood, trichloroethylene and its metabolites in blood or urine, arsenic or methyl parathion metabolites in urine, and PCBs or dioxins in fat tissue. Just as air or ground water monitoring data can provide increased (or decreased) confidence in the results of predictive air or ground water models, biomarkers can be used in a similar manner to evaluate how much confidence should be placed in predictive exposure assessment models. Biological data can be subject to the same shortcomings as other exposure data in terms of data quality and representativeness. The design and performance of the biological data collection effort generally should be carefully evaluated for these factors (e.g., low, medium, and high quality or confidence; low or high bias, etc.) before using the results in the risk decision.

EXHIBIT 4-4

EXAMPLES OF PHYSICAL OR GEOGRAPHICAL FACTORS THAT CAN AFFECT EXPOSURE

- geographical features that limit or enhance accessability (e.g., slopes, valleys, mountains)
- land use, including where exposure occurs within the exposure unit, and the current or future manner in which the receptor contacts contaminated media
- depth of contamination (e.g., surface soil is of greatest concern for direct contact)
- bioavailability of contaminant from media or water (e.g., physicochemical factors that enhance or reduce absorption)
- availability of contaminated medium for exposure (e.g., chipping or peeling paint surfaces vs. intact surfaces)
- water quality and distribution systems, including water hardness and use of lead-soldered pipes
- temporary barriers (e.g., fences, ground cover, and concrete) that affect current (but not necessarily future) exposures

4.4.4 TOXICITY DATA

A variety of factors will affect the magnitude of adverse responses expected to occur in similarly exposed individuals such as age, physiological status, nutritional status, and genotype. In general, these sources of interindividual variability, and related uncertainties, are taken into account in the derivation of toxicity values (e.g., RfC, RfD, CSF) used in human health risk assessments. Thus, human health toxicity values usually are derived to be health-protective for the most sensitive portions of the U.S. population.

- *Sources of variability or uncertainty are typically accounted for in the derivation of toxicity values. Risk managers generally should consider this in making their decision about the appropriate RME risk percentile to use.*

Risk managers should carefully consider whether the toxicity value is representative of the population of concern (Exhibit 4-5). Uncertainty in toxicity values may reflect insufficient data to evaluate developmental toxicity concerns, *in utero* exposures, and chemical mixtures. Also, the population at the site may not be adequately represented by the sensitive part of the U.S. population, accounted for in the derivation of the toxicity value. This determination would require coordination with a toxicologist to review of the basis for the derivation of the toxicity values in question. Even then, in most cases, the determination will be very difficult, because our understanding of human variability in toxicologic responses is very limited. For example, in the derivation of non-cancer human health toxicity values (RfD, RfC), inter-individual variability in toxicokinetics and toxicodynamics is usually represented with an uncertainty factor because data are insufficient to support a more quantitative representation of these sources of inter-individual variability.

EXHIBIT 4-5

EXAMPLES OF TOXICITY CONSIDERATIONS

- How severe is the effect?
- Is the effect reversible?
- How steep is the slope of the dose-response curve at low dose?
- Is the contaminant persistent in the environment or in people?
- Dose the contaminant bioconcentrate as it moves through the food chain?
- How bioavailable is the contaminant?

At one Superfund site, a distribution of toxicity values for methylmercury was used to evaluate exposures to women of child-bearing age. A physiologically-based pharmacokinetic (PBPK) model with distributions for all the model parameters was used to determine a range of mercury levels in hair for a single dose rate. A technique called Benchmark Dose Modeling (BMD) was used to relate the levels in maternal hair to adverse developmental effects, based on data from a prospective epidemiology study. The level of methylmercury in maternal hair thought to present no developmental risk was estimated with the BMD model and used to determine a range of intakes based on the range of physiological parameters in the PBPK model. This range of intakes yielded a Reference Dose for methylmercury expressed as a distribution with a range from 0.0001 to 0.0003 mg/kg-day. While this type of information should not be used to characterize variability in human health risks, it does

provide additional information on uncertainty that can be compared with other sources of uncertainty in the exposure assessment. Thus, toxicity data can inform qualitative risk management choices regarding the percentile representing the RME (within the 90th to 99.9th percentile range) and/or the appropriate confidence limit on the RME estimate.

4.4.5 MULTIPLE CRITERIA FORM THE BASIS OF THE REMEDIAL DECISION

The considerations discussed above may tend to interact in a consistent pattern. Consider, for example, that a risk manager is presented with a risk assessment for a heavy metal in residential soil in which the distribution of risk estimates in the RME range (i.e., 90th to 99.9th percentiles) overlaps the risk range of concern (10^{-6} to 10^{-4}). The risk manager then proceeds with the site technical team to evaluate the data input to the risk assessment, the site-specific factors, and the available biological data. The risk manager learns that the soil collection/analysis effort was well-designed and, hence, has high confidence in the results. However, generic national data were used for the intake and uptake exposure assumptions in the risk assessment. The risk manager also notes that the predominant chemical and physical forms of the metal in the soil are characterized by relatively low bioavailability. In addition, all of the yards in the residential neighborhood are covered with grass lawns, a feature generally expected to reduce direct exposure to soil. The risk manager is also aware of biological data regarding human exposures obtained by the local health department. Concentrations of the metal were measured in samples of human tissue that are indicative of short term and chronic exposures. The levels were within normal physiological ranges, suggesting little, if any, contaminant exposure occurred at the site.

When all of these factors are considered, they form a logical picture, which may suggest that the results of the risk assessment are biased towards an overestimate of risk. These results, together with the nine criteria for remedy selection given by the NCP (U.S. EPA, 1990) (Exhibit 4-6), may form the basis for selecting a cleanup level. A sensitivity analysis should be performed (see Chapter 2) to determine whether the information is relevant for the exposure variables that contribute most to the variability or uncertainty in risk. For example, if absorption fraction contributes little to the variance in a risk estimate, then site-specific data on bioavailability would be less important than information on other variables. Site-specific information can provide greater confidence that a probabilistic risk estimate in the lower end of the RME range (i.e., 90th to 99.9th percentile) is appropriate and health protective.

EXHIBIT 4-6

NINE CRITERIA FOR EVALUATION OF CLEANUP ALTERNATIVES (U.S. EPA, 1990[1])

Threshold Criteria

1. Overall protection of human health and the environment
2. Compliance with ARARs

Balancing Criteria

3. Long-term effectiveness and permanence
4. Reduction in toxicity, mobility, or volume through treatment
5. Short-term effectiveness
6. Implementability
7. Cost

Modifying Criteria

8. State acceptance

4.4.6 EXAMPLE OF RISK ESTIMATES FROM BOTH POINT ESTIMATE AND PRA METHODOLOGIES

The following example, based to large extent on Smith (1994), illustrates the application of point estimate and probabilistic approaches to risk assessment. It is a "bare-bones" example, providing the minimum information that should be presented to a risk manager.

Two chemicals of concern have been identified in groundwater with a point estimate risk assessment: 1,1-dichloroethene and 1,1,2-trichloroethane. The single exposure pathway considered is tap water ingestion. Carcinogenic risks are summed for all contaminants. Table 4-2 summarizes the point estimate and PRA inputs used to estimate variability in lifetime cancer risks for a residential adult population exposed to volatile organic chemicals that have migrated into a residential well. Water samples were obtained from the same well at different points in time over a period of one year. These data characterize temporal variability in the concentrations, which happen to fit a lognormal distribution very well. While one may be tempted to use the lognormal distribution in the 1-D MCA for variability, care should be taken to recognize that the EPC still represents the long-term average concentration. The appropriate metric in this case is the arithmetic mean (or uncertainty in the mean represented by the 95% UCL). Table 4-3 summarizes hypothetical concentration data measured in residential wells. Estimates of the true long-term contaminant concentrations for both CTE and RME individuals are based on the 95% upper confidence limits for the true mean concentration (U.S. EPA, 1992).

The CTE and RME risk from the point estimate approach were calculated using EPA's default exposure assumptions for adults (U.S. EPA, 1993). The PRA risk estimates reflect interindividual variability in exposure. For this example, uncertainty is characterized only by the use of 95% UCL for the mean concentration. For simplicity, correlations between contaminant concentrations, temporal patterns in concentrations, and inhalation exposure during showering were not incorporated into the exposure model. (Generally, a comprehensive PRA would evaluate inhalation exposure due to showering for volatile chemicals in groundwater.) The probability distributions used for other exposure variables reflect a combination of lognormal and triangular distributions. Parameter values are based on values given in the *Exposure Factors Handbook* (U.S. EPA, 1997c). The citations to the primary literature are also provided in Table 4-3. Hypothetical truncation limits were applied to the exposure duration (maximum = 30 years) and the adult body weight (minimum of 25 kg, maximum of 180 kg) to explore this source of uncertainty. These parameter estimates are hypothetical, and constrain the random sampling of the Monte Carlo simulation to plausible bounds (see Chapter 3).

Figure 4-6 shows a risk distribution estimated from a Monte Carlo simulation using 10,000 iterations along with values corresponding to selected percentiles. Both the RME risk from the PRA (determined to be the 95^{th} percentile of the risk distribution for this site) and the RME from the point estimate approach exceed a target risk of $1x10^{-5}$. The risk level of concern for this example corresponds with approximately the 90^{th} percentile (i.e., there is a 10 percent probability that risk exceeds $1x10^{-5}$). The RME range (90^{th} to 99.9^{th} percentiles) spans nearly one order of magnitude. At this point, a sensitivity analysis would be conducted to identify which exposure variables contribute most to the variability in risk. This information may lead to additional probabilistic analysis and/or data collection in order to quantify and reduce sources

of uncertainty in the risk estimates within the RME range. In practice, a sensitivity analysis would be conducted both before and after performing the 1-D MCA.

As stated in Chapter 1, in PRA, a recommended starting point for risk management decisions regarding the RME is the 95th percentile of the risk distribution. The intent of this descriptor is to convey an estimate of risk in the upper end of the risk distribution, while providing a risk manager with some flexibility to select a different (higher or lower) percentile depending on site-specific information regarding exposure and toxicity. Results of the PRA from this example suggest that remedial action may be warranted to reduce risks at the site. The example presented in Section 4.4.7 provides additional guidance on selecting an appropriate percentile from the RME range.

4.4.7 EXAMPLE OF SELECTING AN RME RISK FROM THE RME RANGE

The following example of a hypothetical risk assessment illustrates the type of information that may be considered in determining an appropriate percentile of the RME range (i.e., 90th to 99.9th percentiles) to represent the RME for remedial decisions.

4.4.7.1 BACKGROUND

Site XYZ is a wood treatment industrial facility that has been in operation for twenty years. The owner plans to replace the existing facility with an industrial complex. As part of a recent RI/FS, PAHs from creosote contamination were identified in surface soil. The likely future receptor at the site is an occupational worker (e.g., operator/mover performing light industry, warehousing, or industrial park types of activities). Risk managers selected a cancer risk of $1x10^{-5}$ as a level of concern for the site.

4.4.7.2 RISK ASSESSMENT ASSUMPTIONS

The single chemical of concern was benzo(a)pyrene equivalents, representing carcinogenic PAHs. The exposure scenario is an adult worker who is exposed via soil ingestion. As part of the screening level assessment (Tier 1 of Figure 1-4), point estimates of CTE and RME risk were calculated as $4.0x10^{-6}$ and $8.0x10^{-5}$, based on pilot data for concentrations in soil and default exposure factors. A preliminary sensitivity analysis demonstrated that the risk estimates are most sensitive to the following two exposure variables – concentration in soil and exposure duration.

Table 4-2. Inputs for Drinking Water Exposure Scenario

Input Variable			Point Estimate		Probability Distributions and Point Estimates		
Symbol	Description	Units	CTE	RME	Type	Parameters	Source
C_w	concentration in water	µg/L	95% UCL	95% UCL	95% UCL	see Table 4-2	hypothetical site data
IR_w	tap water ingestion rate, adults ages 20 - 65 yrs	mL/day	1400	2000	lognormal[1]	(7.023, 0.489)	EFH (U.S. EPA, 1997a); Roseberry & Burmaster, 1992
EF	exposure frequency	days/yr	234	350	triangular[3]	(180, 250, 350)	judgment
ED	exposure duration	yrs	9	30	truncated lognormal[2]	(4.2, 5.0, 0, 30)	site-specific survey
BW	body weight, adults	kg	70	70	truncated lognormal[2]	(71, 15.9, 30, 150)	Brainard & Burmaster, 1992
AT	averaging time	days	25550	25550	constant	25550	RAGS (U.S. EPA, 1989)
CSF	oral cancer slope factor	$(mg/kg\text{-}day)^{-1}$	1,1 - Dichloroethene = 6.00E-01 1,1,2 - Trichloroethane = 5.72E-02				IRIS, June 1998

[1] Parameters of lognormal PDF are X ~ exp[Normal{AM of ln(X), SD of ln(X)}].

[2] Parameters of lognormal PDF are X ~ Log(AM of X, SD of X, min, max)

[3] Parameters of triangular PDF are X ~ Triang{min, best estimate, max}.

Table 4-3. Concentration Data for Chemicals of Concern

Contaminant	Cw (µg/L)[1]		
	AM	SD	95% UCL
1,1-Dichloroethene	1.0	1.8	2.2
1,1,2-Trichloroethane	1.2	2.2	2.5

[1]Concentrations were fit to a lognormal distribution; 95% UCL was calculated using the Land method (U.S. EPA, 1992).
AM = arithmetic mean

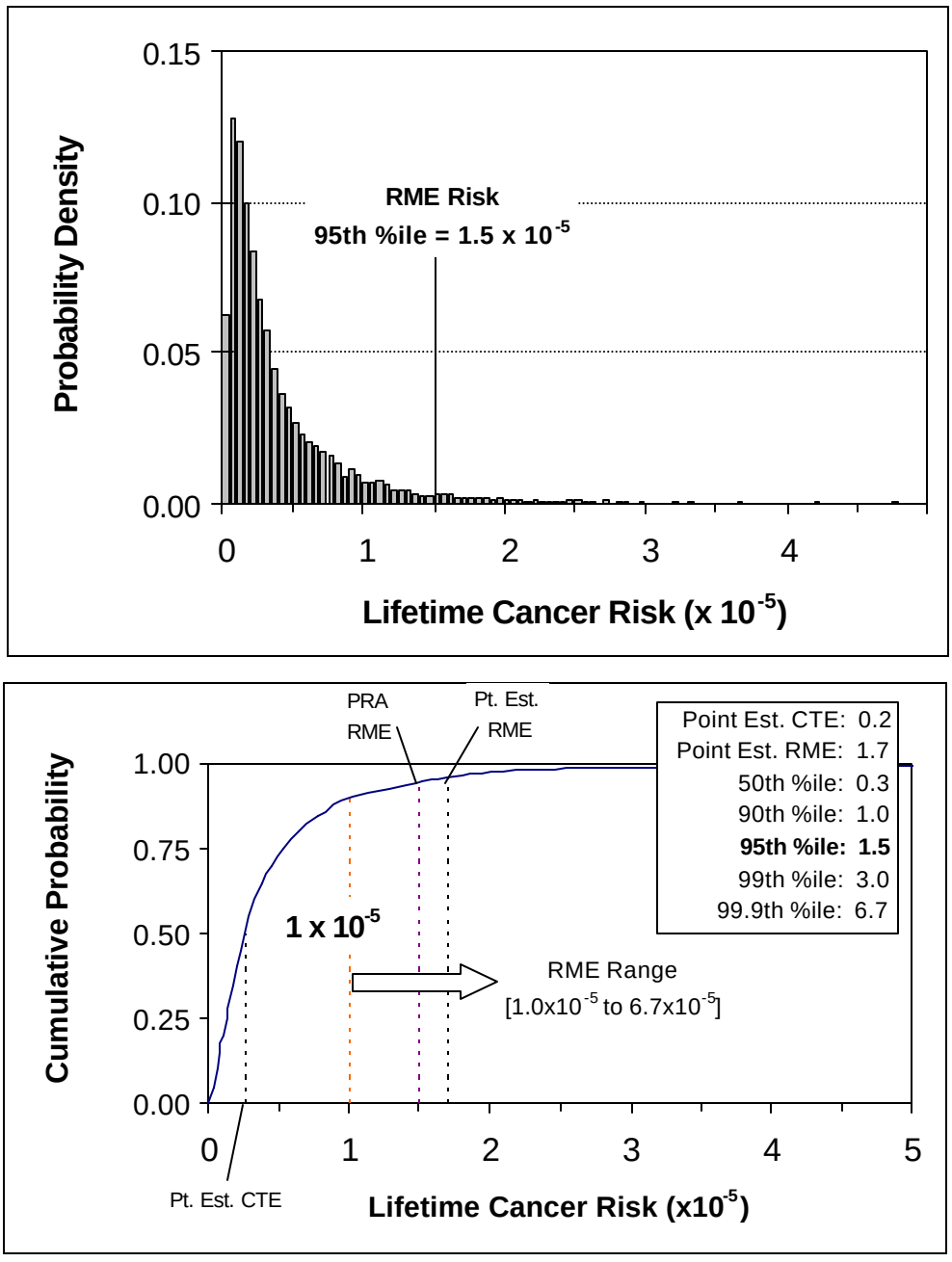


Figure 4-6. 1-D Monte Carlo simulation results (n = 10,000 iterations) showing the PDF (top panel) and CDF (bottom panel) for the distribution of risk based on a hypothetical exposure scenario. Important summary information is presented for both the point estimate approach (CTE and RME) and the PRA. The risk criterion, selected as $1x10^{-5}$, falls at the low end of the RME range (90^{th} to 99.9^{th} percentiles).

4.4.7.3 PROBABILISTIC RISK ASSESSMENT

The site owner expressed a wish to perform a PRA. A scoping meeting was held with the EPA remedial project manager and branch chief, regional risk assessors, the site owner and his consultant, state regulators, and representatives from two local community groups. The stakeholders agreed that the data in hand was insufficient to support a PRA and recommended collecting additional soil samples and site-specific estimates of job tenure. The stakeholders also decided that a 1-D MCA of variability in risk would provide useful information for the remedial decision.

Soil samples were analyzed with an immunological analysis for PAHs that could be performed onsite. Twenty percent of the immunological results were confirmed with a CLP methods. A regression equation that related B(a)P equivalents to the immunologic results was developed based on the 20% of the soil samples that were jointly analyzed. Use of the immunologic analysis provided a cost-effective way to obtain many more soil measurements.

All workers at the plant completed a confidential questionnaire regarding the length of time each had worked at Site XYZ and how long each planned to remain employed there. The median job tenure at Site XYZ was 1.8 years, considerably less than the 6.6 year national average.

For this scenario, assume the results of a sensitivity analysis suggested that the high-end risk estimates are greatly affected by the variability in exposure duration (i.e., job tenure). An RME point estimate of 5.0 $x10^{-5}$ was developed using the 95th percentile value of the job tenure distribution as the exposure duration and the 95% UCL based on all the soil sampling. A PRA using 1-D MCA was performed using the site-specific empirical distribution of job tenure and the 95% UCL as a point estimate for concentration. Table 4-4 summarizes the risk estimates from a Monte Carlo simulation with 10,000 iterations.

Table 4-4. Selected risk estimates from 1-D MCA simulation.

Percentile	Risk (x 10^{-5})
90.0	2.1
92.0	2.4
95.0	2.9
97.0	3.5
99.0	5.1
99.9	9.8

4.4.7.4 RME RISK FROM THE RME RANGE

When determining cleanup levels that are protective of the RME individual, the remedial project manager's task is to select a percentile from the RME range of 90 to 99.9 that represents the RME risk estimate. This guidance recommends the 95th percentile as a reasonable starting point. As discussed in Sections 1.4 and 4.4.1, the selection of a different percentile will require judgment regarding the quantity and quality of site-specific information. Factors that tend to mitigate risk may support the choice of a lower percentile in the RME range. Factors that may tend to exacerbate risk support the choice of moving to higher than the 95th percentile. Uncertainties that cannot be adequately quantified by the chosen modeling approach may also support the selection of a higher percentile value to estimate the RME. In practice, a risk manager must balance multiple, competing factors.

At Site XYZ, the remedial project manager considered four factors:

Factors that mitigate risk or reduce uncertainty

1) Variability in exposure duration was characterized from reliable site-specific data. Job tenure at Site XYZ was considerably lower than initially thought. Since this was the most sensitive variable in the model, the remedial project manager has confidence in the high-end risk estimates.

Factors that increase uncertainty in the risk distribution

2) Industrial activities would involve periodic movement of heavy machinery within exposed (bare) soil areas. Without knowing the patterns of moving equipment, it is reasonable to assume that individuals exposed more often to bare soil areas experience higher soil ingestion rates, and potentially different exposure point concentrations. Thus, some individuals may be highly exposed.

3) A regression equation used to relate B(a)P equivalents to cPAHs constituted a source of parameter uncertainty in the concentration term that was not reflected by the use of the 95% UCL.

4) Adult soil ingestion was characterized by a point estimate rather than a probability distribution, which limited the information that could be obtained from the sensitivity analysis following the 1-D MCA. Qualitative information regarding activity patterns of adult workers suggested that contact-intensive activities may yield higher ingestion rates than represented by even the RME point estimate.

Based on these considerations, the remedial project manager decided to select the 97^{th} percentile for the RME. The reliable site-specific information on exposure duration was offset by the sources of uncertainty that were not easily quantified in the exposure model.

Uncertainty in the three factors discussed above was addressed qualitatively in both the PRA and point estimate approaches. Selection of the 97^{th} percentile, rather than the 95^{th} percentile, did not change the result of the risk assessment – the risk estimate at both these percentiles exceeded $1x10^{-5}$. However, the decision to select a percentile toward the upper end of the RME range does have implications for determining the appropriate PRG. The selection of a PRG based on the assumptions in this example is discussed further in Chapter 7.

REFERENCES FOR CHAPTER 4

Brainard, J. and D. E. Burmaster. 1992. Bivariate distributions for height and weight of men and women in the United States. *Risk Anal*. 12(2): 267-275.

Burmaster, D.E. and P.D. Anderson. 1994. Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessment. *Risk Anal.* 14(4): 477-481.

Morgan, G.M. and M. Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge University Press, New York.

Roseberry, A.M. and D.E. Burmaster. 1992. Lognormal distributions for water intake by children and adults. *Risk Anal.* 12(1): 99-104.

Smith, R.L. 1994. Uses of Monte Carlo simulation for human exposure assessment at a Superfund site. *Risk Anal*. 14(4): 433-439.

U.S. EPA. 1989. *Risk Assessment Guidance for Superfund (RAGS): Volume I. Human Health Evaluation Manual (HHEM) (Part A, Baseline Risk Assessment)*. Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002. NTIS PB90-155581.

U.S. EPA. 1990. National Oil and Hazardous Substances Pollution Contingency Plan. Final Rule. 40 CFR 300: *55 Federal Register*, 8666-8865, March 8.

U.S. EPA. 1991. *Role of the Baseline Risk Assessment in Superfund Remedy Selection Decisions*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive 9355.0-30.

U.S. EPA. 1992. *Supplemental Guidance to RAGS: Calculating the Concentration Term*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive 9285.7-081.

U.S. EPA. 1993a. *Superfund's Standard Default Exposure Factors for the Central Tendency and Reasonable Maximum Exposure*. Draft. Office of Emergency and Remedial Response, Washington, DC. November.

U.S. EPA. 1993b. *Data Quality Objectives Process for Superfund*. Office of Solid Waste and Emergency Response, Washington, DC. EPA

U.S. EPA. 1995a. *Memorandum from Carol Browner on Policy on Evaluating Health Risks to Children*. Office of the Administrator, Washington, DC. October 20, 1995.

U.S. EPA. 1995b. *Memorandum from Carol Browner on Risk Characterization*. Office of the Administrator, Washington, DC. February 22, 1995.

U.S. EPA. 1996. *Memorandum from Carol Browner on EPA's Report, Environmental Health Threats to Children*. Office of the Administrator, Washington, DC. September 1996.

U.S. EPA. 1997a. *The Lognormal Distribution in Environmental Applications.* Office of Research and Development & Office of Solid Waste and Emergency Response, Washington, DC. EPA/600/R-97/006.

U.S. EPA. 1997b. *Memorandum from Deputy Administrator Fred Hansen on the Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment*, and *Guiding Principles for Monte Carlo Analysis.* Office of Research and Development, Washington, DC. EPA/630/R-97/001. May 15, 1997.

U.S. EPA. 1997c. *Exposure Factors Handbook*. Office of Research and Development, Washington, DC. EPA/600/P-95/002Fa, Fb, and Fc.

U.S. EPA. 1998. *Integrated Risk Information System*. Online. Office of Research and Development, National Center for Environmental Assessment, Cincinnati, OH.

CHAPTER 5
USING PROBABILISTIC ANALYSIS IN ECOLOGICAL RISK ASSESSMENT

5.0 INTRODUCTION

Basic Approach for Performing Ecological Risk Assessments

Ecological risk assessment is a key component of the Remedial Investigation process that EPA uses at many Superfund sites. EPA has developed extensive guidance and policies on methods and approaches for performing ecological risk assessments, including the following:

1. *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments – Interim Final ("ERAGS")*, (U.S. EPA, 1997a). This document includes processes and steps specifically selected for use in ecological risk assessments at Superfund sites. This document supersedes the 1989 *EPA RAGS, Vol. 2, Environmental Evaluation Manual*. Supplements to RAGS Vol. 2 and ERAGS include the *EcoUpdates* (U.S. EPA, OSWER, Intermittent Bulletin Series, 1991 to 1994), which provide brief recommendations on common issues for Superfund ecological risk assessments.

2. *Ecological Risk Assessment and Risk Management Principles for Superfund Sites* (U.S. EPA, 1999). This document supplements ERAGS (U.S. EPA, 1997a) and provides additional guidance for both risk assessment and risk management.

3. *Guidelines for Ecological Risk Assessment ("Guidelines")* (U.S. EPA, 1998a). This document updates general (non-program specific) guidance that expands upon and replaces the earlier *Framework for Ecological Risk Assessment* (U.S. EPA, 1992a).

4. *Risk Assessment Guidance for Superfund (RAGS): Volume 1- Human Health Evaluation Manual (Part D, Standardized Planning, Reporting, and Review of Superfund Risk Assessments, Interim Final)*, (U.S. EPA, 1998b). This guidance specifies formats that are recommended to present data and results in baseline risk assessments (both human and ecological) at Superfund sites.

5. Policy Memorandum: *Guidance on Risk Characterization for Risk Managers and Risk Assessors*, F. Henry Habicht, Deputy Administrator, Feb. 26, 1992 (U.S. EPA, 1992b). This policy recommends baseline risk assessments to present ranges of risks based on "central tendency" and "reasonable maximum" (RME) or "high end" exposures with corresponding risk estimates.

6. Policy Memorandum*: Role of the Ecological Risk Assessment in the Baseline Risk Assessment* from Elliott Laws, Assistant Administrator, August 12, 1994, OSWER Directive No. 9285.7-17 (U.S. EPA, 1994). This policy recommends the same high level of effort and quality for ecological risk assessments as commonly performed for human health risk assessments.

7. Policy Memorandum: *EPA Risk Characterization Program from Carol Browner, Administrator, March 21, 1995* (U.S. EPA, 1995a). This policy clarifies the presentation of hazards and uncertainty in human and ecological risk assessments, calling for clarity, transparency, reasonableness, and consistency.

EXHIBIT 5-1
DEFINITIONS FOR CHAPTER 5

Assessment endpoint - An explicit expression of the environmental feature that is to be protected, operationally defined for risk assessment as valuable attributes of an ecological entity.

Community - An assemblage of populations of different species specified by locales in space and time.

Conceptual model - A site conceptual model (SCM) in the problem formulation for an ecological risk assessment is a written description and visual representation of predicted relationships between ecological entities and the stressors to which they may be exposed, including sources and pathways of stressors.

Ecological risk assessment (ERA) - The process that evaluates the likelihood that adverse ecological effects may occur or are occurring as a result of exposure to one or more stressors.

Lines of evidence - Information derived from different sources or techniques that can be used to characterize the level of risk posed to exposed receptors; weight of evidence generally refers to the quantity of science, while strength of evidence generally refers to the quality of science.

Lowest observed-adverse-effect-level (LOAEL) - The lowest level of a stressor evaluated in a test (when administered to an ecological receptor) that causes biological and/or statistically significant differences in responses (adverse effects) from the controls.

Measurement endpoint - A measurable ecological property that is related to the valued characteristic chosen as the assessment endpoint. Measurement endpoints often are expressed as the statistical or numeric summaries of the observations that make up the measurement.

No observed adverse effect level (NOAEL) - The highest level of a stressor administered to an ecological receptor in a test that does not cause biological and/or statistically significant differences in responses (adverse effects) from controls.

Population - An aggregate of individuals of a species within a specified location in space and time.

Receptor - The ecological entity (with various levels of organization) exposed to the stressor.

Risk characterization (ecological) - The third and last phase of ERA that integrates the analyses of exposure to stressors with associated ecological effects to evaluate likelihoods of adverse ecological effects. The ecological relevance of the adverse effects is discussed, including consideration of the types, severity, and magnitudes of the effects, their spatial and temporal patterns, and the likelihood of recovery.

Scientific/Management Decision Point (SMDP) - A time during the ERA when a risk assessor communicates results or plans at that stage to a risk manager. The risk manager decides if information is sufficient to proceed with risk management strategies or whether more information is needed to characterize risk.

Species - A group of organisms that actually or potentially interbreed and are reproductively isolated from similar groups; also, a taxonomic grouping of morphologically similar individuals.

Stressor - Any chemical, physical, or biological entity that can induce an adverse response in an ecological receptor; Superfund focuses on chemical (toxicant) stressors.

Toxicity Reference Value (TRV) - A risk-based dose or concentration that usually includes factors of toxic uncertainties and is often based on a NOAEL or LOAEL; a TRV is sometimes referred to as a "toxicity benchmark", but this is not the same as the conventional "benchmark dose", which is the lower probability bound on a dose for a designated low response.

1

Ecological risk assessment (ERA) is defined by the 1998 EPA *Guidelines* as an evaluation of the "likelihood that adverse ecological effects may occur or are occurring as a result of exposure to more stressors" (U.S. EPA, 1998a). Figure 5-1 summarizes the 3-phase paradigm for performing an ecological risk assessment as recommended by the *Guidelines*. The three phases include problem formulation, analysis of exposure and effects, and risk characterization; these phases generally follow the four steps of the paradigm for risk assessment and management that were described and refined by the National Academy of Sciences (NAS, 1983; 1994).

- **Problem Formulation** (Phase 1) provides a foundation for the entire risk assessment and includes the development of a site conceptual model with exposure pathways, the specification of risk management goals, the selection of assessment endpoints, and the development of a sampling and analysis plan to collect data on measurement endpoints that are needed to support the ecological risk assessment. In general, problem formulation is an iterative process, and substantial re-evaluation may occur as new information and data are collected during the site investigation evaluation. Collection of data in subsequent iterations is often triggered by identification of large uncertainties in the risk characterization that prohibit confident risk management decision making.

- **Analysis** (Phase 2) includes two principal steps: characterization of exposure and characterization of ecological effects. Exposure characterization describes sources of stressors, their distribution in the environment, and their contact or co-occurrence with ecological receptors. Ecological effects characterization evaluates stressor-response relationships or evidence that exposure to stressors causes, or has the potential to cause, an observed response. This potential for effect is usually summarized in a Toxicity Reference Value (TRV) that describes some measure of exposure or dose that is or is not associated with the occurrence of an adverse ecological effect. The exposure characterization and the ecological effects characterization are then combined in the risk characterization.

- **Risk Characterization** (Phase 3) estimates the nature and severity of ecological risks attributable to exposure to stressors at a site, and interprets the relevance of the adversity of ecological effects. Policies by EPA recommend that a range of risks be presented (U.S. EPA, 1992b), that certain formats for results be employed (U.S. EPA, 1997a, 1998b), and that ERAs impart clarity, transparency, reasonableness, and consistency (U.S. EPA, 1995a). Good risk characterizations will express results clearly, articulate major assumptions and uncertainties, identify reasonable alternative interpretations, and separate the scientific conclusions from policy judgments (U.S. EPA, 1998a). In addition, risk characterizations generally should be as quantitative as possible, identify acceptably safe exposure levels to stressors (i.e., PRGs), and identify major data gaps that might require subsequent iterations of data collection or analysis.

Specific guidance for completion of ecological risk assessments within the Superfund Program is provided in the *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments* (ERAGS) (U.S. EPA, 1997a), supplemented by *Ecological Risk Assessment and Risk Management Principles for Superfund Sites* (U.S. EPA, 1999). In Superfund risk assessment, the focus is always on chemical stressors that have been released into the environment. The *ERAGS* document describes an eight-step process for completing an ecological risk

assessment for Superfund, as shown in Figure 5-2. Although separated into somewhat different organizational steps, this process is generally similar to, and consistent with, the approach described in the 1998 EPA *Guidelines*.

Figure 5-1. The general framework for ecological risk assessment recommended by U.S. EPA 1998.

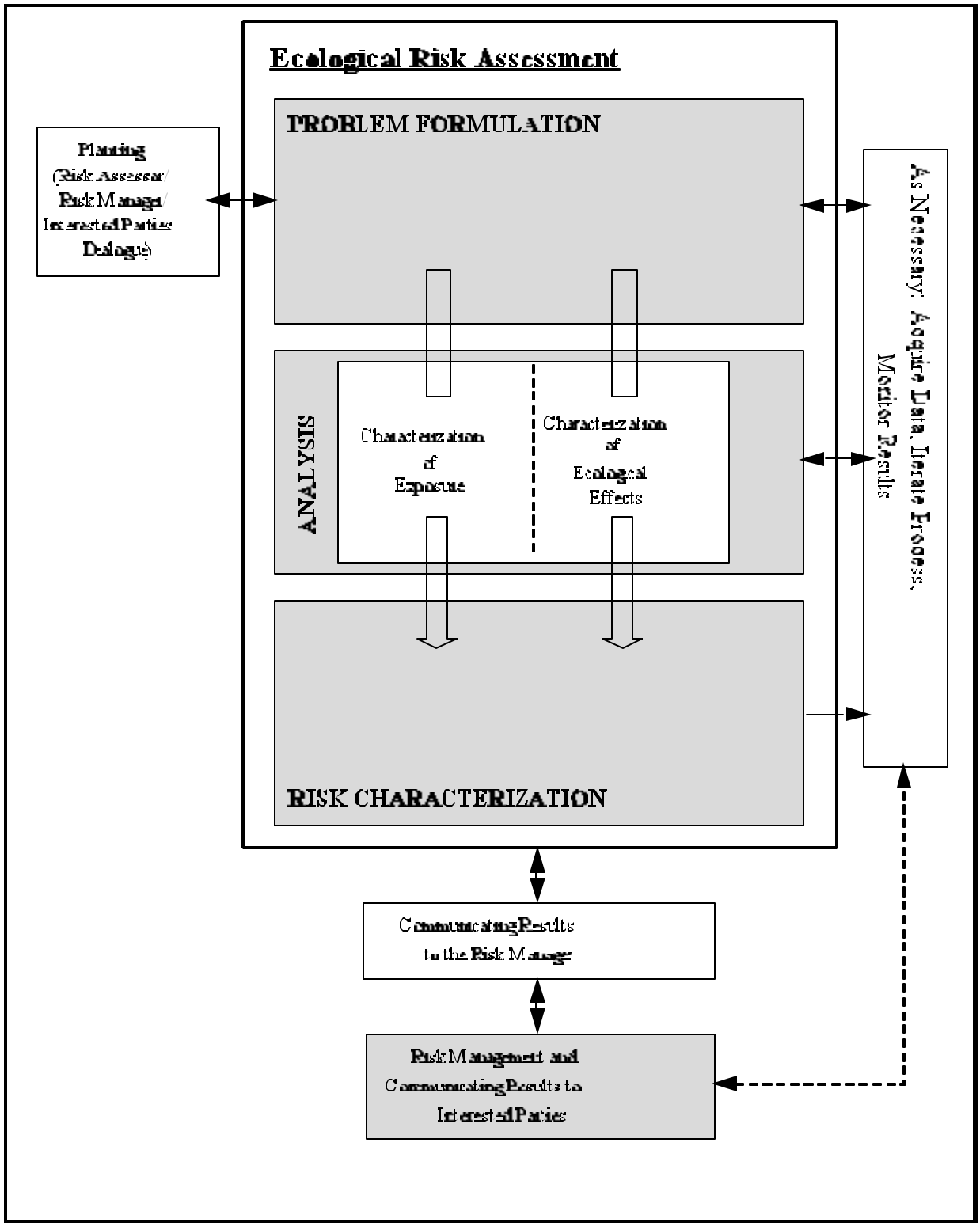
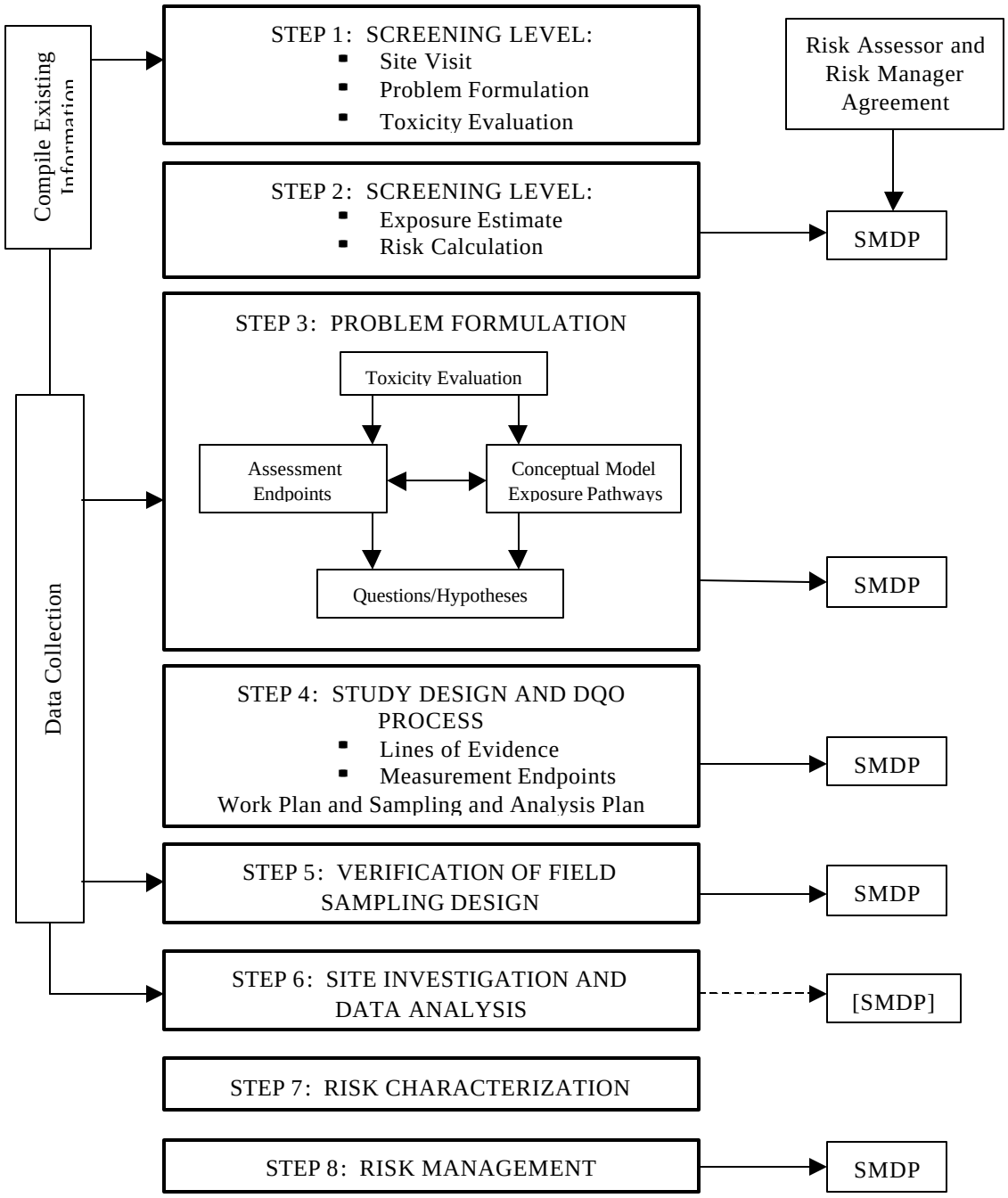


Figure 5-2. The eight-step ecological risk assessment process recommended for evaluation of Superfund sites (U.S. EPA, 1997a). SMDP = scientific/management decision point.



Focus of this Chapter

This chapter focuses on the use of probabilistic risk assessment (PRA) for evaluation of ecological exposures and risks at Superfund sites. The basic mathematical methods and approaches used in performing a PRA for a human health risk assessment and an ecological risk assessment are essentially the same. It is assumed that readers of this chapter are already familiar with the basic concepts and principles of ecological risk assessment, as detailed in the Agency guidance and policy documents cited above, and with the basic principles of PRA that are described in other chapters of this document.

- *The basic mathematical principles and techniques of PRA are generally the same for ecological and human health risk assessment.*

The four main topics covered in this chapter are summarized below:

- Section 5.1: understanding the basics of PRA in ecological risk assessment, with attention to key differences between the use of PRA in human and ecological risk assessments;
- Section 5.2: using a decision process for determining if and when to use a PRA in ERAs;
- Section 5.3: developing presentations of ecological risk estimates from PRA for decision making, and communicating ecological PRA results with risk managers and the public; and
- Section 5.4: general guidelines for submitting a PRA as part of an ERA.

5.1 PRA IN ECOLOGICAL RISK ASSESSMENT

A flow chart that illustrates how the PRA process may be incorporated into an ecological risk assessment is shown in Figure 5-3. This figure is similar in content and concept to that shown previously for human health risk assessment (see Figure 1-4). As seen in Figure 5-3, performance of a PRA should not normally be considered until the screening ERA and an initial iteration of a point estimate risk analysis are both complete, and when the results of these analyses indicate that performance of a PRA will likely provide additional useful information to the risk manager. If a PRA is judged to be of potential use to the risk manager, three alternative (and not mutually exclusive) types of PRAs may be considered, as follows:

1. Variability between individuals in a population with regard to the level of exposure (dose), toxicity, and/or risk (1-dimensional MCA (1-D MCA) of variability). This option, which is frequently used in human health risk assessment, is also valuable in many ecological risk assessments, especially where there may be risks to threatened or endangered species (where risks to individuals are more important), and when the assessment of population risk requires knowledge of the frequency or severity of risks to different individuals within the population.

2. Uncertainty in the central tendency (average) level of exposure (dose), toxicity, and/or risk (1-D MCA of uncertainty) to a population or community of receptors. This option is useful in cases when risks to an average receptor (rather than an RME receptor) are judged to be most relevant to risk management decision making. Uncertainty around the mean (CTE) estimate of exposure

or risk can be characterized using one-dimensional modeling, while uncertainty around other statistics of the variability distribution generally requires 2-dimensional modeling (see below).

3. Variability and uncertainty in the exposure (dose), toxicity, and/or risk (2-dimensional MCA (2-D MCA) of both variability and uncertainty) to a population or community of receptors. This option is used for human health risk to estimate the best estimate and the uncertainty range around the RME, and the same approach can help describe the full range of risks (including both CTE and RME) and the associated uncertainty for ecological receptors.

EXHIBIT 5-2

DIFFERENCES IN PRA APPLIED TO ECOLOGICAL AND HUMAN HEALTH RISK ASSESSMENT

- data needs and data availability
- endpoints
- toxicity evaluation

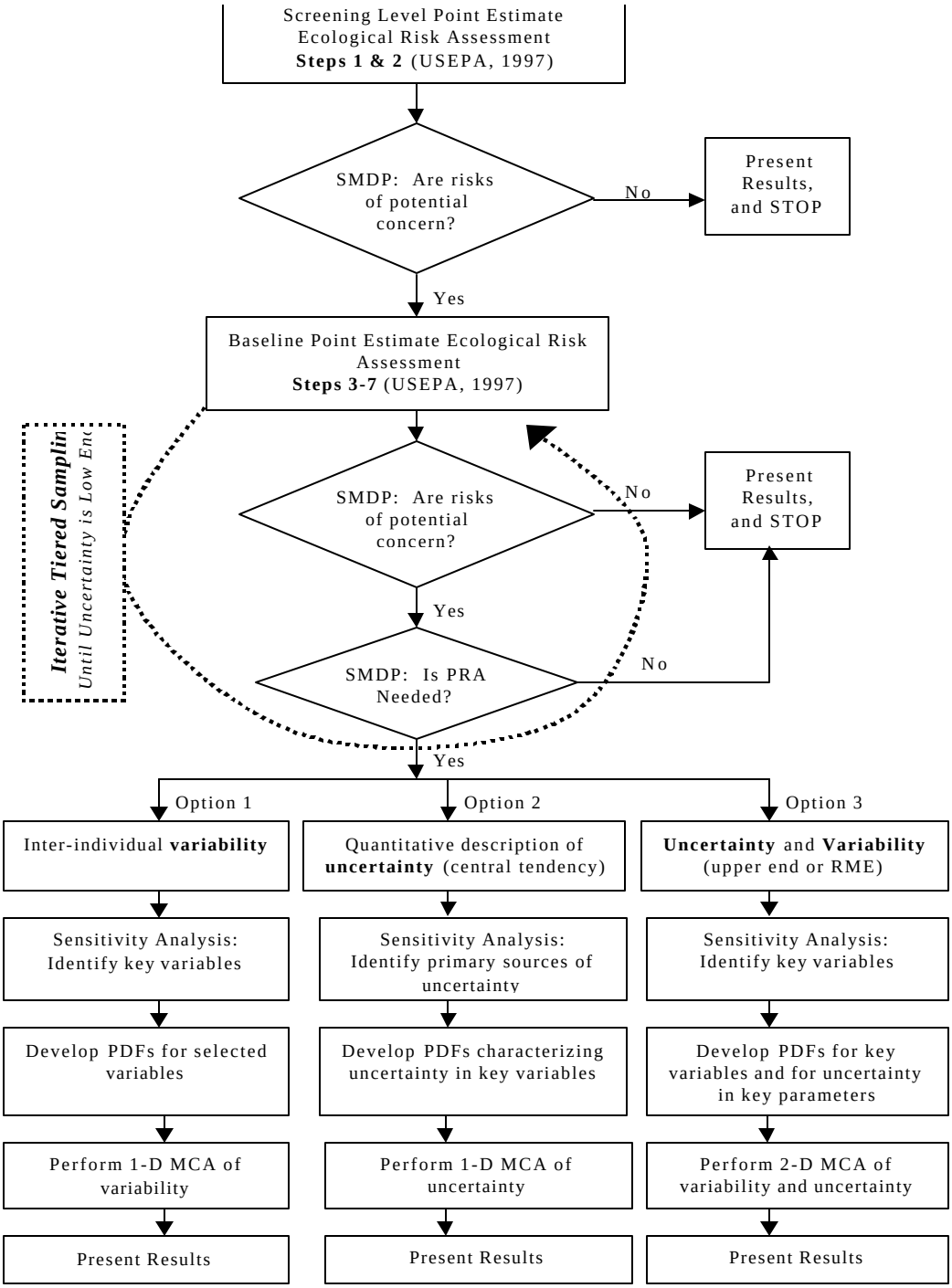
Several important technical and tactical differences between human health and ecological risk assessments that may influence how PRA is used in ecological assessments are noted in Exhibit 5-2 and are discussed below.

5.1.1 DATA NEEDS AND AVAILABILITY FOR VARIABILITY ANALYSIS

Due to their complexity, ecological risk assessments are often more demanding of time and resources than are human health risk assessments. This complexity can occur at many different levels of the analysis. For example, ecological risk assessments often evaluate multiple receptors, require complex food web measurements or modeling, require sophisticated methods for defining exposure units and deriving exposure point concentrations, and account for complex behaviors of some species.

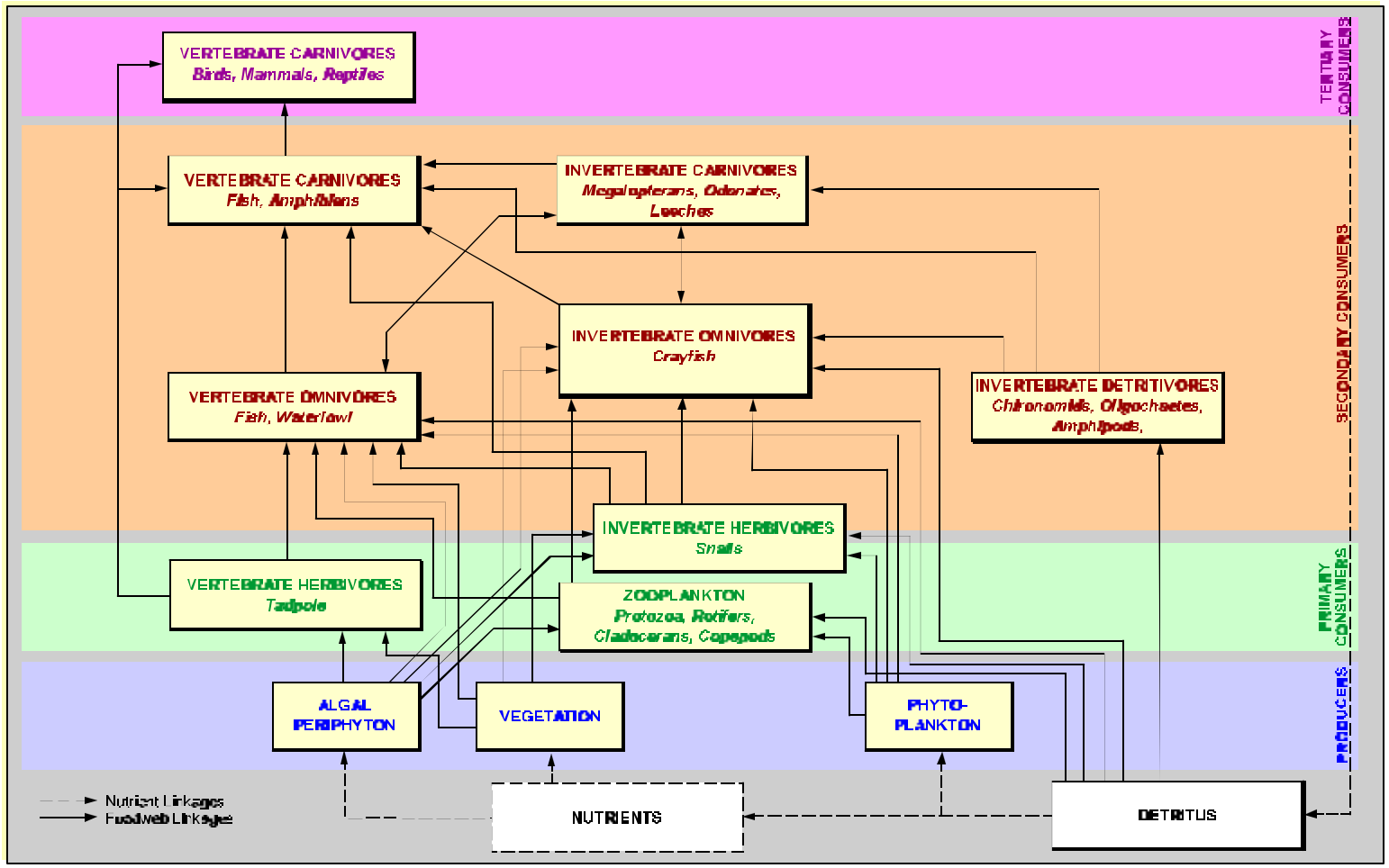
An example of a food-web that is part of a site conceptual model illustrating this complexity is shown in Figure 5-4. Focusing on vertebrate omnivores, this example shows that eight different food-web pathways potentially contribute to the exposure of receptors in this class. In addition, exposure of these receptors to site chemicals may also occur from ingestion of water, soil, and/or sediment (not shown). Assuming that at least two variables are needed to quantify exposure by each pathway, and that some pathways (e.g., food-web pathways) may require 3-5 variables, it is clear that complete point estimate evaluation of this class of receptors may include numeric values for as many as 25-50 variables. In order to perform a comprehensive PRA, assuming that each variable is described by at least two parameters to define its PDF, a total in excess of 50-100 parameters may be needed. A similar effort would also be expected in order to evaluate each of the other groups of receptors selected for quantitative evaluation.

Figure 5-3. Potential applications of probabilistic risk assessment in ecological assessments.



SDMP = Scientific/Management Decision Point

Figure 5-4. Example of a portion of a site conceptual model (SCM) that shows pathways of food-web exposure for ecological receptors in an estuarine habitat.



One of the challenges in characterizing variability in ecological risk assessments is the disparity between data needs and data sources. For example, while EPA does provide a summary of wildlife exposure factors (U.S. EPA, 1993a), there are no standard default exposure factors for ERAs analogous to those that have been developed by EPA for human risk assessments. Furthermore, while a number of distributions useful for evaluation of human exposures have been published in the literature, there are very few published estimates of distributions available for use in ERA at this time. Thus, the effort to derive PDFs (either from published data or by performing site-specific field studies) for all variables can be substantial and potentially prohibitive. In general, the costs and benefits of performing an ecological PRA should be carefully evaluated before beginning a PRA (see Section 5.2).

Fortunately, the list of variables and parameters that may be important sources of variability or uncertainty in an ecological PRA can usually be reduced by screening the exposure pathways in the site conceptual model in order to identify the "risk drivers". As discussed in Chapter 2, a sensitivity analysis can be used to identify the subset of exposure pathways and exposure media that dominate the risks to a population of ecological receptors. To a lesser extent, variables and parameters that contribute greatly to variability or uncertainty in the exposure and risk model may be characterized using the point estimate approach. Variables and parameters that contribute minimally to the output distribution should generally be described by point estimates. Because sensitivity analysis can help to focus a PRA on the sources of variability and uncertainty that may affect a risk management decision, it is an important initial step in the tiered process described above. In general, the greatest effort should be devoted to deriving PDFs for those key variables or key uncertain parameters that contribute most to the variability and/or uncertainty in risk.

5.1.2 ENDPOINTS

According to EPA policy (U.S. EPA, 1992b), both human and ecological risk assessments should present variability in risks, generally by describing the range of risks that apply to different parts of the variability distribution, including both the CTE and "upper end" RME (see Exhibit 5-3). For risks to human health, risk management decision-making normally focuses on the *RME individual* who is at the upper end of the inter-individual distribution of exposures and risk. In general, the risk management objective is to ensure that risks to the RME individual do not exceed some specified level of protection. In contrast, many ecological risk assessments focus on local *population sustainability and community integrity*. Because of this general difference in the risk management objective, ecological risks to a population are sometimes evaluated by assessing the risks to an average (rather than an RME) individual in the population.

The significance of this difference is that, if risk to a population is to be assessed by evaluation of the CTE receptor, it is not always necessary to use PRA techniques to characterize the full range of inter-individual variability in exposure and risk. Rather, it may be more helpful to use PRA techniques to characterize the *uncertainty* in the estimate of average or central tendency (see Option 2 in Figure 5-3). In this special case, uncertainty can be characterized using 1-D MCA rather than 2-D MCA. This type of uncertainty analysis can be especially helpful for presenting ranges of risks to managers, since it provides a much clearer picture of the confidence associated with a particular point estimate of risk. Quantitative estimates of uncertainty can provide risk managers with greater latitude in selecting appropriate remedies to protect the health (in terms of the sustainability and integrity) of ecological

1
2

EXHIBIT 5-3

RISK DESCRIPTORS

As described in Chapter 1, the EPA Superfund program generally focuses on two types or risk descriptors: *high-end* risks and *central tendency* risks.

High-end Risk Estimate

"The high end risk descriptor is a plausible estimate of the individual risk for those persons at the upper end of the risk distribution. The intent of this descriptor is to convey an estimate of risk in the upper range of the distribution, but to avoid estimates, which are beyond the true distribution" (U.S. EPA, 1992b). This term conveys the upper range of the distribution (e.g., above the 90th percentile of the risk distribution) (U.S. EPA, 1992b, 1995a).

High-end risk is typically based on an estimate of the reasonable maximum exposure (RME), defined as the highest individual exposure that is reasonably expected to occur at a site over time (U.S. EPA, 1989a, 1989b, 1990, 1992a).

Discussions on high-end risk descriptors in other chapters within this document focus on describing the "RME" individual. This term should be interpreted as the more general "high-end risk" for purposes of ecological risk assessment. The endpoint of concern for ecological risk assessment is typically local population-level effects (e.g., population sustainability and community integrity), rather than individual-level effects (EPA, 1998a).

Central Tendency Risk Estimate

By contrast, the central tendency risk descriptor represents a central level of risk, referred to as "average" or "typical" risk. Likewise, central tendency risk is based on the central tendency exposure (CTE).

Presented together, the high-end and central tendency risks provide a measure of both the magnitude as well as the expected variability in risks among the exposed population. High-end risk descriptors (e.g., RME) are typically most relevant when the risk management objective is the protection of human health at the individual level. Central tendency risk descriptors may be more relevant for ecological risk assessments that focus on population-level risks.

populations. In addition, a 1-D MCA for uncertainty can help indicate whether additional tiers of sampling may reduce uncertainties in risk estimates.

When the risk to a receptor (either at the individual level or the population level) cannot be properly characterized by consideration of CTE risks alone, then 2-D MCA may be needed to simulate both the inter-individual variability distribution and to characterize the uncertainty in that distribution (see Option 3 in Figure 5-3). For example, 2-D MCA is generally needed to quantify uncertainty in RME risk estimates at the upper tail of the risk distribution, especially if the receptor is a federal or state-listed threatened or endangered species.

EXHIBIT 5-4

EXAMPLES OF OUTPUT FROM PRA

- distribution of risk among individuals in an exposed population
- probability of a response occurring either in an individual or in a population
- uncertainty in a point estimate of risk
- uncertainty in the site conceptual exposure model
- primary sources of variability and/or uncertainty (sensitivity analysis)
- probability that exposure exceeds a specific benchmark dose or concentration
- association(s) between site media concentrations and the probability of an adverse effect

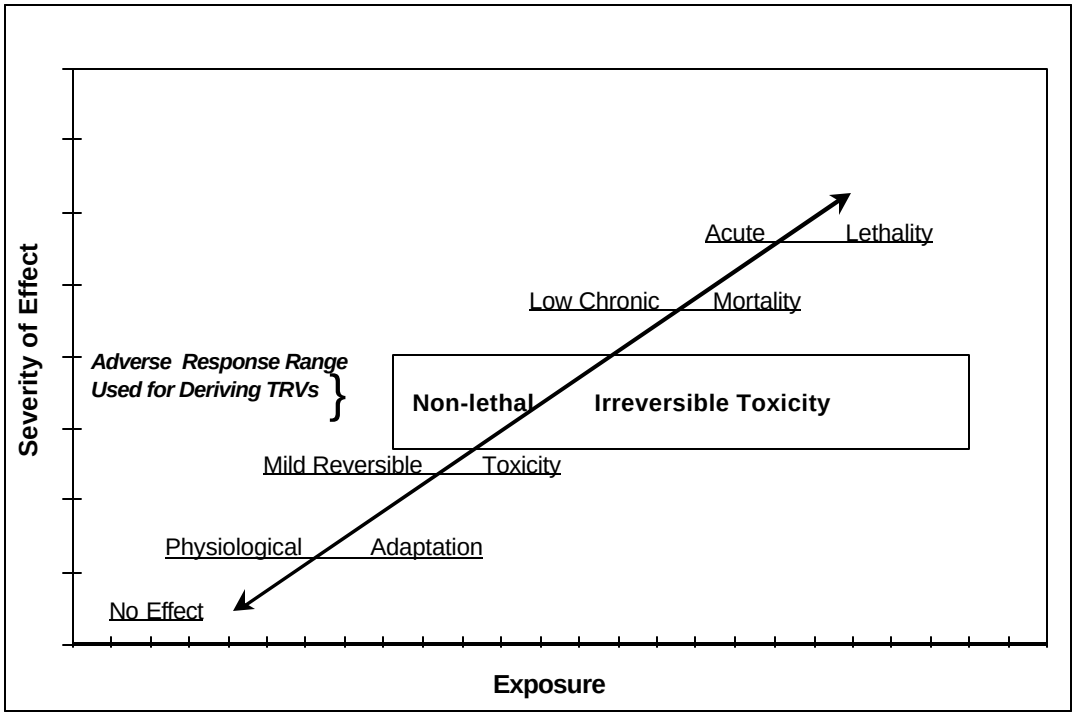
5.1.3 TOXICITY EVALUATION

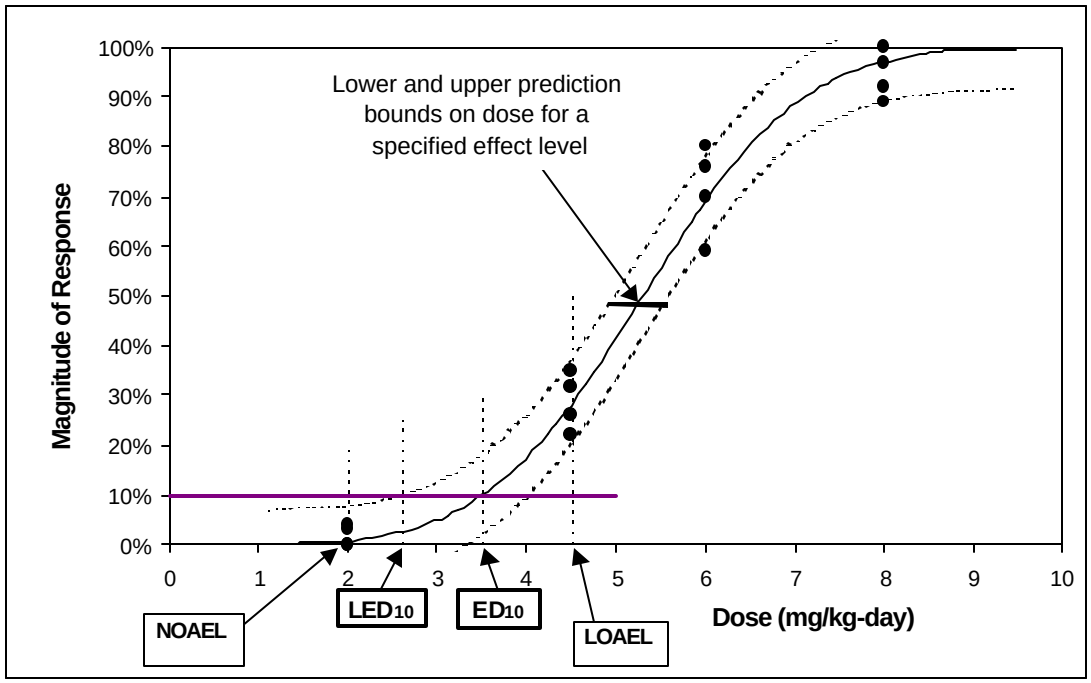
The toxicity reference value (TRV) that is used in ecological risk assessments is similar in concept to the Reference Dose (RfD) used in human health risk assessments. Generally, an ecological TRV is based on a non-lethal, population-relevant, adverse effect. That is, mild, reversible changes are usually not selected as the basis of the TRV because the occurrence of this type of effect is not usually a good predictor of ecologically significant individual or population-level adverse effects. Conversely, selection of overly severe adverse effects (including acute or chronic lethality) is generally not appropriate, since a TRV based on such a severe effect may not protect organisms from the occurrence of less severe (but still significant) adverse effects. This concept is illustrated in Figure 5-5 (upper panel).

In ecological risk assessment, a TRV is often presented as a range of two values that are based on the experimentally observed NOAEL and the LOAEL for a particular endpoint. Alternatively, TRVs may be derived by Benchmark Dose (BMD) modeling (U.S. EPA, 1995b), which uses available dose-response data to estimate the dose which produces a 10% effect (ED10), and the lower 95% confidence bound on the ED10 (the LED10). The LED10 is often used as the BMD. Figure 5-5 (bottom panel)

illustrates the different approaches for estimating one or more TRVs from a hypothetical dose-response data set.

Figure 5-5. (Top) Selection of toxicological endpoints for derivation of TRVs based on severity of effect. (Bottom) Example toxicological dose-response curve based on the magnitude of a response in individual animals, showing the NOAEL, LOAEL, ED10 and LED10. The LED10 is often used as a Benchmark Dose (BMD).

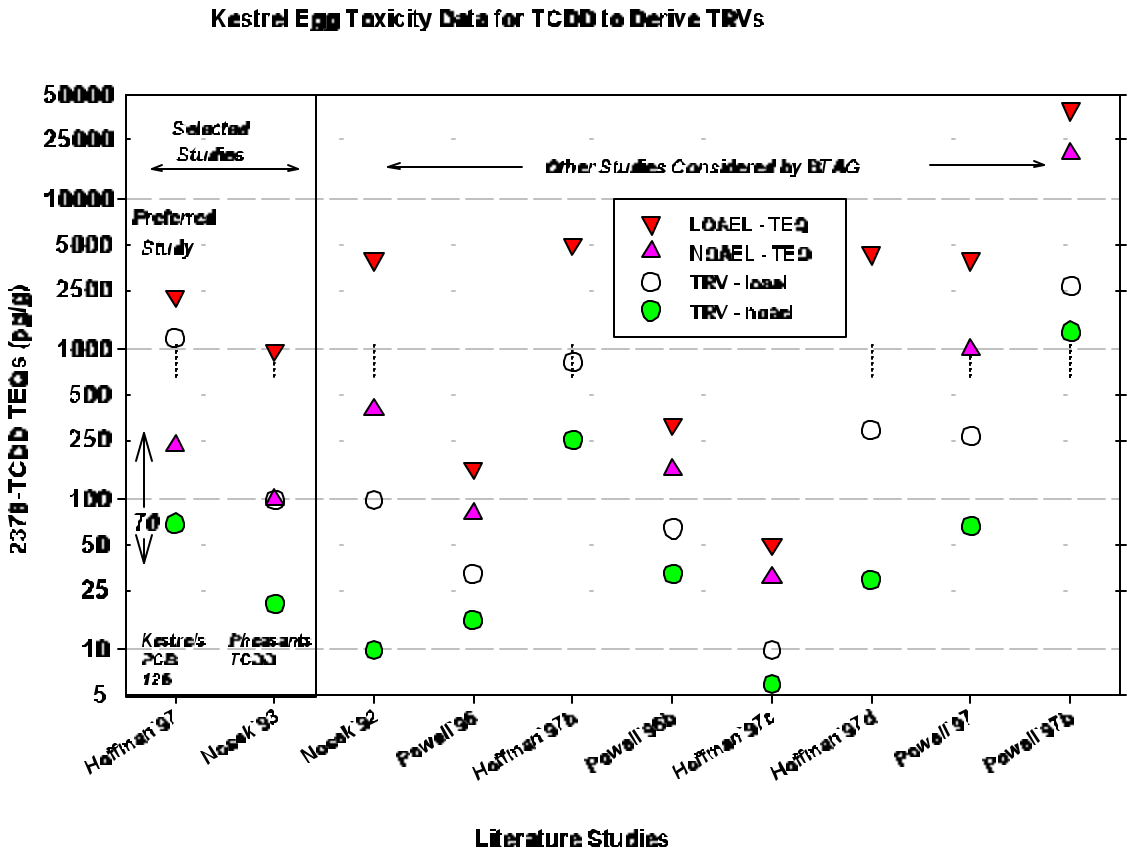




Choosing which data set to employ in deriving a TRV is often difficult. Figure 5-6 illustrates a case where multiple studies were performed on the toxicity of 1,2,7,8-TCDD to kestrel eggs. As shown in this example, typically there is significant variation in effect-levels between different studies. A careful evaluation of the relative strengths and limitations of each study can generally help to identify the strongest "key studies" that most appropriately support the derivation of a TRV.

FIGURE 5-6. SELECTION OF TRVS FOR ECOLOGICAL RISK ASSESSMENT

From a toxicological perspective, TRVs (NOAELs, LOAELs, ED_x) generally should be extrapolated from toxicological studies that have the least uncertainty and most applicability with regard to: species tested, duration of chronic exposure, severity of endpoint, and other scientific factors related to good experimental design and conduct. Sometimes it is difficult to identify a "clearly best" toxicity dose-response study and results from which to derive valid TRVs. Uncertainty analysis can be used to sample from the better candidate study endpoints. The following is an example of the presentation of extrapolated dose-response values from individual study NOAELs and LOAELs to derive distributions for uncertainty in TRV.



In this example, chronic TRVs were derived from a variety of literature studies that reported NOAELs and LOAELs from dioxin-like exposures in American kestrels. The varied, but normalized, results for the TRV-loaels (lowest circles) are presented together to illustrate that these values can be used to create an uncertainty distribution with upper and lower bounds surrounding a most-likely value (mode); e.g., a
1 triangular function with 70 ppt being the most likely value with bounds of about 6 to 1200 ppt.

In ecological risk assessments, any specific TRV (NOAEL, LOAEL, BMD, etc) may be characterized by a probability distribution that defines either variability (e.g., inter-individual variability in toxicity) or uncertainty (e.g., uncertainty in the best point estimate of the mean TRV). By contrast, current EPA policy for human health PRAs is generally to characterize the RfD by a point estimate, usually based on a single key study (see Chapter 1).

If a TRV is to be evaluated as a variable, the raw toxicity data (i.e., the response for each animal from studies that expose multiple animals to a range of doses) can be used to characterize inter-individual variability in response, both at a fixed dose and across different doses. A variety of PDFs (e.g., normal, lognormal, beta, triangular) may also be appropriate for characterizing inter-individual variability in response. Methods for selecting and fitting distributions to data are discussed in Chapter 3.

For example, consider the hypothetical dose-response data set below:

Dose (∶g/kg-d)	Response (% incidence)		
	Mean	Stdev	CV
0	2.9	1.0	0.34
4	6.2	1.9	0.31
8	14	4.5	0.32
12	33	10	0.30
20	78	18	0.23
30	98	21	0.21
40	107	28	0.26

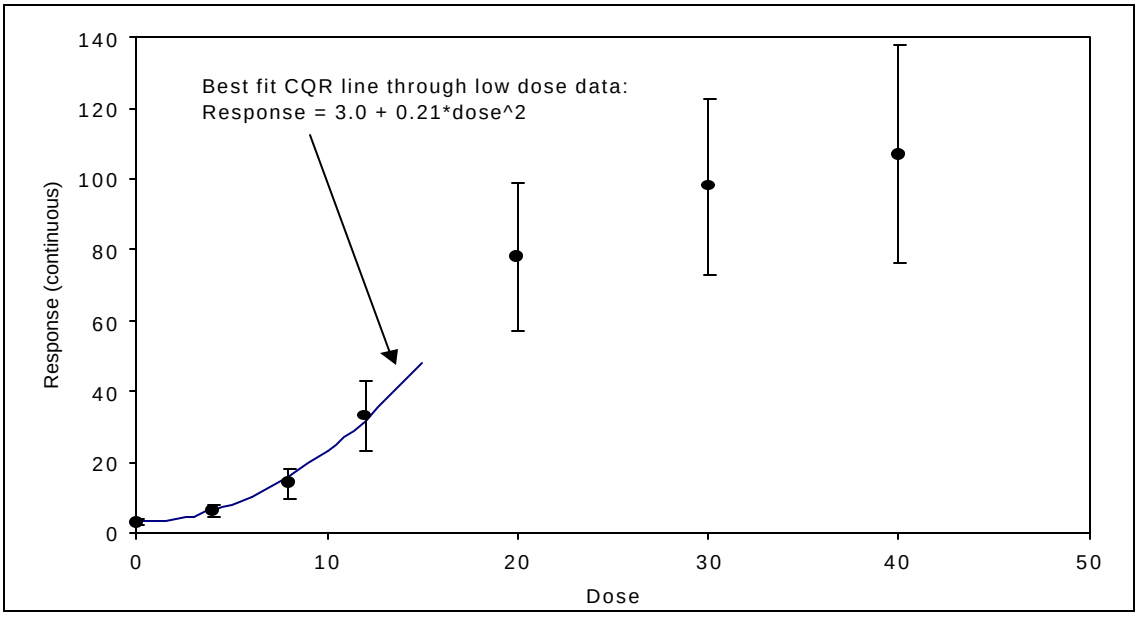
Figure 5-7. (Upper panel) plots these summary statistics (mean, standard deviation), and shows the best-fit continuous quadratic regression (CQR) equation line through the mean responses of the low dose groups. Because (in this example) the raw data are not provided, it is not possible to determine the distribution of individual responses at a given dose. One approach for describing this source of inter-individual variability is to assume the parameter estimates (AM, SD) define a truncated normal distribution of the following form:

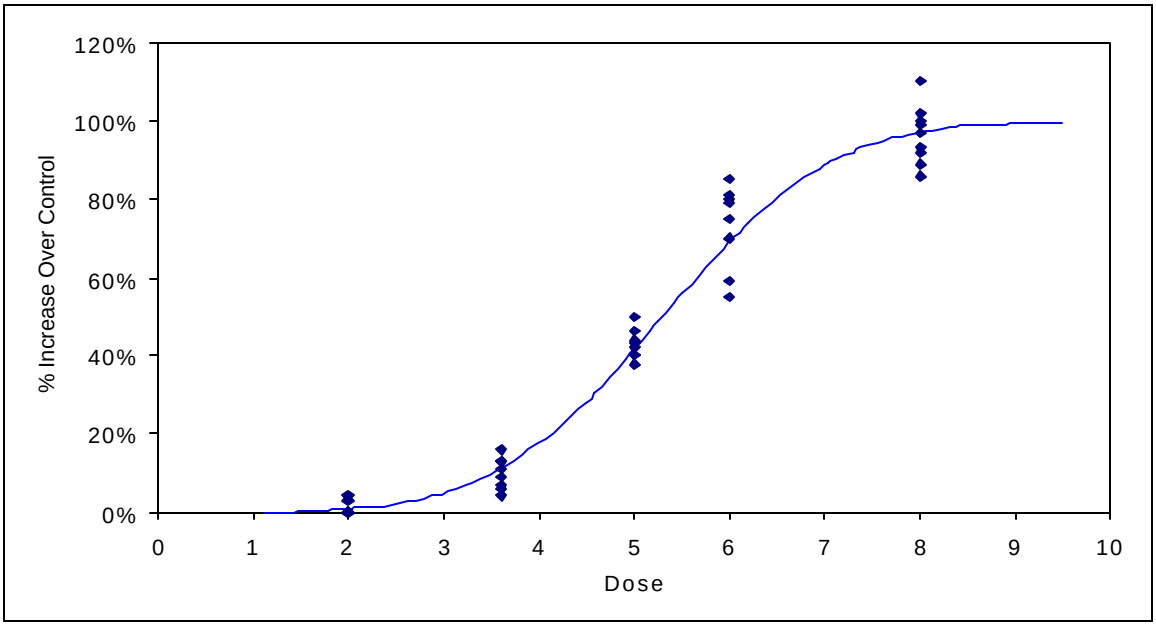
$$Response_d \sim Truncated\ Normal\ (m_d, s_d, min, max)$$

where,

m_d = mean response at dose "d"
s_d = standard deviation of responses at dose "d"
min = minimum plausible response
max = maximum plausible response

Figure 5-7. (Top) Hypothetical dose-response data used to describe inter-individual variability in response to a chemical stressor. The solid circle is the mean response at each dose tested, and the error bars (the standard deviation) reflects the variability in response between different animals. The line through the low dose data is the best fit Continuous Quadratic Regression (CQR) line. The line is not extrapolated to higher doses since this type of equation does not plateau. (Bottom) Example dose-response data set with data for individual animals. If the data are sufficient, the data set within a dose group can be used to model inter-individual variability in response.





A truncated distribution is used to prevent selection of implausible values from an un-bounded PDF. Judgment is needed to estimate the bounds (truncation limits) of the distribution, but a conventional choice such as plus or minus two times the standard deviation may be appropriate. Because (in this example) the coefficient of variation (CV = SD/AM) is approximately constant (about 0.3), the standard deviation (and hence the bounds of the distribution) can be estimated from the mean response at each dose. Thus, each of the parameters of the truncated normal distribution of inter-individual variability in response might be modeled as a function of dose as follows:

$$m_d = 3.0 + 0.21\ (dose)^2$$
$$s_d = 0.3\ m_d$$
$$min_d = m_d - 2\ s_d$$
$$max_d = m_d + 2\ s_d$$

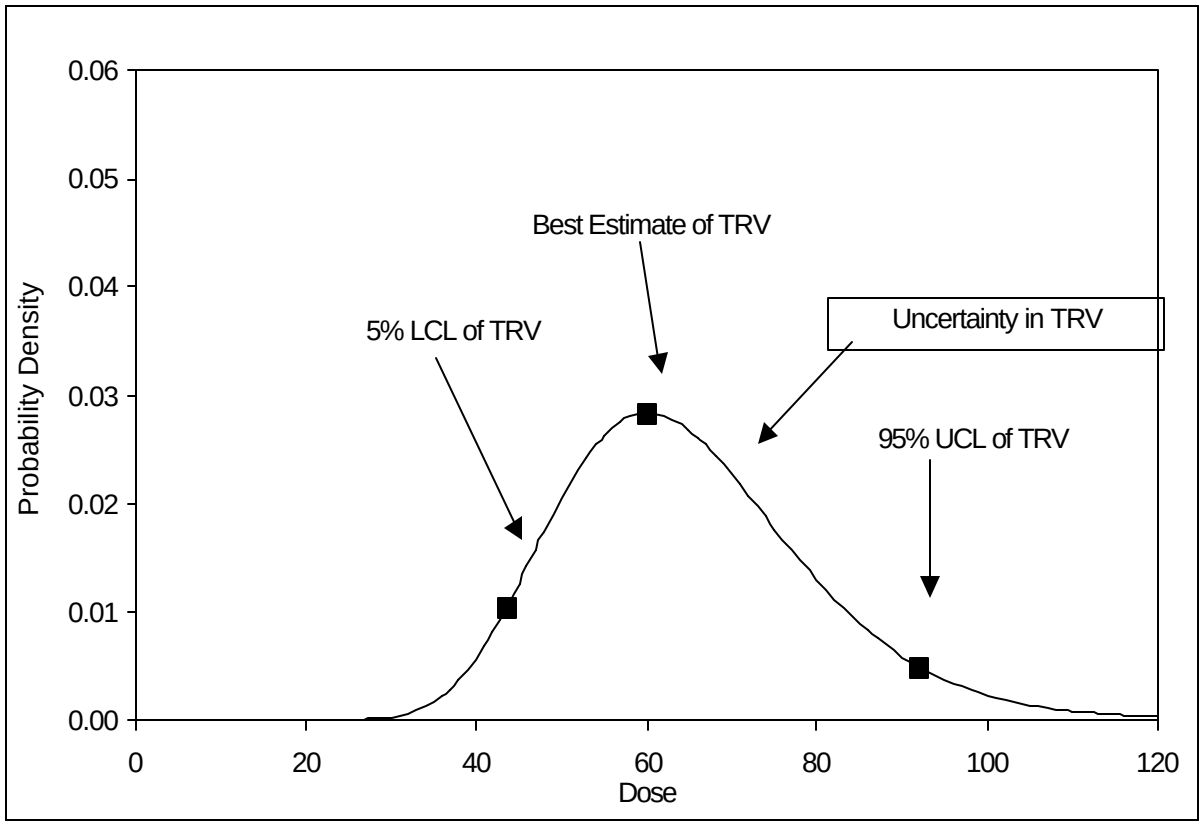
Figure 5-7 (bottom panel) illustrates an example where the individual data points are available from a selected toxicological study. If sufficient numbers of animals are available, an empirical distribution can be fit to the data within each dose group using a commercial distribution-fitting program, and these distributions can then be used to predict the range of responses in a group of receptors exposed at some specified level. In cases where the number of animals per group is too small to permit rigorous curve fitting, a distribution of responses may be modeled based on the statistics (mean, standard deviation, minimum, maximum) of the responses within a group, similar to the approach described above.

When the number of animals is too small to support any meaningful evaluation of inter-individual variability in response within a dose group (e.g., data are available only for 2-3 animals per group), it is still often helpful to quantify uncertainty in the point estimate of TRV (i.e., TRV_U). For example, the best estimate of the TRV that is extrapolated from a dose-response study might be some value "x", but the true value might lie in the range [a, b]. Thus, TRV_U might be modeled as a simple triangular PDF:

TRV_u - Triangular (a,x,b)

Alternative probability distributions for modeling uncertainty may be fit to the data, or estimated using bootstrap techniques (Chapter 3, Appendix C). Figure 5-8 illustrates a lognormal PDF for TRV_U. The uncertainty distribution for a TRV can sometimes be based on an evaluation of the differences between studies that have measured the same endpoint (e.g., see Figure 5-6), or on an evaluation of the uncertainty within a study (e.g., a range might span the lower bound to the upper bound of the ED10). In general, the distribution chosen to model uncertainty in a TRV should reflect all relevant data that are available (both within and between studies).

Figure 5-8. Example PDF reflecting uncertainty in the true value of a specific toxicity reference value (TRV) such as the NOAEL or the LOAEL for a particular endpoint.



5.2 DECIDING WHEN TO USE PRA IN ECOLOGICAL RISK ASSESSMENT

As shown previously in Figure 5-2, the ecological risk assessment process for Superfund includes a number of scientific/management decision points (SMDPs) (U.S. EPA, 1997a). The SMDP is a point of consultation between the risk manager, the risk assessment team, and other stakeholders, and is intended to provide an opportunity for re-evaluation of direction and goals of the assessment at critical points in the process. It is at these points in the ecological risk assessment that discussions occur concerning uncertainty in the risk assessment, and where the possible utility of PRA should be considered.

One of the key SMDPs is an evaluation of whether or not a PRA is needed. This SMDP is usually considered only after completion of both the screening level and initial iterations of baseline point estimate ecological risk assessments (Steps 1 to 7; U.S. EPA, 1997a; 1999).

- *Results of a point estimate approach generally should be presented together with a PRA.*

The results of a point estimate risk assessment will normally present the range of risks based on CTE and RME input assumptions and on the NOAEL- and LOAEL-based TRVs (U.S. EPA, 1992b and 1997a). These results are generally presented in a 2 x 2 matrix, as shown below. The highest estimate of risk (the largest HQ) is generally derived from the ratio of the RME exposure over the NOAEL-based TRV, while the lowest estimate of risk is generally derived from the ratio of the CTE exposure to the LOAEL-based TRV. These two extremes can be used to screen out cases where PRA is not likely to be useful. That is, if the risk to the RME receptor is below a level of concern using the NOAEL-based TRV, then risks to individuals and the population are likely to be low and PRA analysis is likely not needed. If risks to the CTE receptor are above a level of concern using the LOAEL-based TRV, then risks to individuals and the population are likely to be of concern, and a PRA may or may not provide any additional useful information to the risk manager. If the risks are intermediate between these two bounds, then risks are close to the decision threshold, and PRA evaluation is likely to be helpful in characterizing risks and supporting decision-making.

TRV	**Exposure**	
	CTE	**RME**
NOAEL	Intermediate HQ; use professional judgment to assess need for further evaluation	Highest HQ value; screen out if less than 1.0
LOAEL	Lowest HQ value; screen in if greater than 1	Intermediate HQ; use professional judgment to assess need for further evaluation

When deciding whether or not to use PRA, the type of information that may be useful for risk management should be considered. Exhibit 5-4 gives examples of the information that may be obtained using PRA.

An appropriate method for PRA (see Figure 5-3) can be selected based on the assessment endpoints and goals of the ecological risk assessment. An example of the type of goal that would support each of the three methods is presented below:

1. If a description of the inter-individual variability in exposure and risk is needed, then sensitivity analysis is generally conducted to identify key variables and a 1-D MCA of variability is completed.

2. If a quantitative description of the uncertainty in the *central tendency* risk is desired, then a 1-D MCA can be performed based on the uncertainty around the mean values of the input variables.

3. If the ecological risk assessment requires characterization of the full range of exposures across different individuals in a population, then a 2-D MCA of variability and uncertainty in exposure and effects can be completed to quantify the uncertainty around the *upper end* as well as the *central tendency* exposure and risk. This type of analysis is usually needed when evaluating a rare, threatened or endangered species, and may be needed for other receptors as well.

This additional information, however, sometimes comes with an additional effort, primarily in the form of added time and resources to collect the necessary data, perform the analyses, and document the methods and the results. Therefore, the risk assessor and risk manager should generally weigh the advantages and disadvantages of the proposed PRA (see Chapter 1). General questions that can be addressed to ascertain if the PRA would provide useful information are highlighted below.

Is there a Clear Benefit to a PRA?

In general, if the point estimate risk assessment results are sufficient to support remedial decision-making (i.e., it is clear either that there is a risk, or that there is no risk) in respect to where and how much to remediate, then a PRA is not recommended. However, if the point estimate risk results lie close to the decision threshold (which is always the case for the boundaries of the gradient of contamination that lie between excess risk and adequate safety), then a PRA is often beneficial. The evaluation of the utility of PRA is, however, expected to vary from site to site based on site-specific conditions and considerations, such as spatial scale. It is for this reason that universal rules for choosing when to do or not to do a PRA of variability and/or uncertainty cannot be made here. Some questions that may help guide the decision process are given below.

- Would additional information from a PRA be useful to improve the basis for, and confidence of, decisions concerning the remedial action?

- What are the additional costs and time associated with the PRA, including any necessary data collection as well as technical review of PRA calculations, results and assumptions?

- Are the potential costs associated with the remediation based on the point estimate results low compared with the costs associated with collecting the additional data needed to perform a useful PRA?

- Would reducing uncertainty help in the selection of the most appropriate remedial alternative?

- Can the degree of confidence (or lack thereof) in point estimates of exposure and risk be adequately described and presented without PRA?

- Is it necessary to rank exposures, exposure pathways, sites or contaminants with estimates of confidence?

- Is it necessary to associate a dose (or concentration) range with a probability of response? Will a point estimate be insufficient for decision making due to excess uncertainty?

Are there Sufficient Data to Complete a PRA of Variability?

Data gaps may be a limiting factor in using PRA for characterizing variability in risk. Often data are lacking for some of the variables in the risk model. If these data gaps are minor and the associated variables do not contribute greatly to risk, a credible PRA may be performed without additional data collection. An initial sensitivity analysis is important for assessing whether data gaps introduce important sources of uncertainty in the risk estimates. Additional data collection efforts generally should be focused on the most influential variables.

Data quality requirements for a PRA of variability are generally similar to those for the point estimate ecological risk assessment, except that emphasis is placed on defining the full probability distribution of key variables rather than just selected statistics (e.g., mean, 95th percentile). These data quality requirements are outlined in the *Guidance for Data Usability in Risk Assessment* (U.S. EPA, 1992c) and in *Data Quality Objectives* (U.S. EPA, 1993b). Sensitivity analyses can help show where additional data can most effectively reduce uncertainty in PRAs.

Will an Analysis of Uncertainty Add Helpful Information?

PRA evaluations of uncertainty (as opposed to variability) are often very helpful in revealing the degree of certainty (or lack of certainty) around any particular estimate of exposure or risk (e.g., the CTE or RME). Quantitative evaluations of uncertainty generally can be performed with little or no additional data acquisition. That is, the purpose of the analysis is to estimate the uncertainty around an exposure or risk estimate *given the data at hand*. The only additional information needed to perform the analysis is an estimate of the uncertainty in the true parameter values of the key variables in the variability model. In some cases, these estimates of uncertainty around parameter values may be developed from statistical analysis of the available data. An example of how this might be done is provided in Exhibit 5-5. Alternatively, professional judgment may be used to establish credible bounds on the parameters, especially when data are sparse.

- *Even in the presence of data gaps, uncertainty analysis using PRA can provide useful information. Indeed, it is precisely when data are limiting or absent that a quantitative analysis of uncertainty may be most helpful.*

EXHIBIT 5-5

EXAMPLE DERIVATION OF UNCERTAINTY BOUNDS FOR THE PARAMETERS OF A PDF

There are many alternative methods that can be used to develop uncertainty bounds around the parameters of a PDF, and the best approach will depend on the nature and extent of data available. Consider the following hypothetical case where the variable of interest is the dietary intake by a field mouse. Three separate studies on intake were located in the literature:

Study	N	Dietary Intake (kg/kg-day)	
		Mean	Stdev
1	10	0.22	0.08
2	5	0.31	0.11
3	20	0.18	0.07

Given that the these three studies are not statistically different from each other, the best estimates of the mean and standard deviation can be calculated by pooling the values across the three studies:

$$\text{Pooled mean} = \sum(N_i * M_i) \,/\, \sum N_i = 0.21$$

$$\text{Pooled stdev} = \text{sqrt}[\sum(N_i-1)*(\text{stdev}_i)^2 \,/\, (\sum N_i - 3)] = 0.079$$

The uncertainty around each of these pooled statistics is difficult to calculate rigorously, since the bounds depend on the shape of the underlying distribution and the number of samples drawn. However, a reasonable approximation can usually be achieved by parametric bootstrapping, or by assuming that the sampling distribution of the pooled mean is t-distributed and the pooled variance is inverse chi-squared distributed (this will be true for pooled sample sizes that are reasonably large). In this case, the uncertainty around the pooled mean and pooled variance can be modeled as:

$$\text{Pooled mean} = 0.21 - (T_{n-1})*0.079/\text{sqrt}(35)$$

$$\text{Pooled stdev} = \text{sqrt}\,[34*(0.079)^2 / \text{CHISQ}_{n-1}]$$

where:

T_{n-1} = T-distribution with n-1 degrees of freedom
CHISQ_{n-1} = Chi-squared distribution with n-1 degrees of freedom

Thus, lack of knowledge about the true parameters of a distribution is generally a weak reason for not performing an uncertainty analysis.

What Are the Key Variables or Parameters? (Sensitivity Analysis)

A quantitative sensitivity analysis (see Chapter 2) is used to identify the key sources of variability and/or uncertainty. Sensitivity analyses can be performed at multiple points in the development of a PRA (see Chapter 1). Sensitivity analysis is particularly useful prior to implementing a preliminary PRA. The results of the sensitivity analyses can be used to focus data collection and modeling efforts in order to best reduce uncertainty and characterize variability.

What are the opinions of the EPA Regional Ecotoxicologist or BTAG ?

Ecological risk assessments for Superfund are usually produced under the direction of an EPA Regional Biological Technical Assistance Group (BTAG) coordinator (U.S. EPA, 1993a) or an EPA regional Ecotoxicologist (U.S. EPA, 1997a). This guidance strongly recommends consulting with the BTAG coordinator at EPA and/or the EPA Regional ecotoxicologist regarding the need and the feasibility of applying a PRA for an ecological risk assessment.

5.3 PRESENTING AND INTERPRETING PROBABILISTIC ECOLOGICAL RISK ASSESSMENTS

There are many alternative methods that can be used to present the results of ecological PRA calculations of variability and/or uncertainty to risk managers and other readers. This section provides a number of examples that illustrate some basic options, and provides a discussion of how the results may be used to derive a better characterization of exposure and risk than is usually possible using a point estimate approach. It is important to note that, whether a point estimate approach is employed alone, or a probabilistic approach is employed in conjunction with the point estimate approach, interpreting the results of an ecological risk assessment cannot be reduced to a single or simple default rule. Rather, the results are generally interpreted based on an understanding of the risk management goals and the assessment endpoint, the toxicological basis of the TRV, and the characteristics of the receptor being assessed.

5.3.1 PRESENTING DESCRIPTIONS OF VARIABILITY

There are several basic options for presenting a description of variability among different individuals within an exposed population, including a) the distribution of exposures (expressed either as dose or as concentration), b) the distribution of HQ and/or HI values (either by chemical and pathway, or summed across chemicals and pathways), and c) the distribution of responses (either the incidence and/or severity of an effect). In all cases, the basic information which this presentation allows is an estimate of the fraction of a population that is exposed above some specified level of exposure or response, which in turn allows for a characterization of risks to individuals and/or the population. Some examples are presented below.

Comparing PRA and Point Estimate Results

In PRA, the results of the point-estimate calculations generally should be presented with the PRA results in order to highlight potential differences in assumptions and confidence in the risk estimates. Specifically, in plots of exposure (dose or concentration), the range from the CTE to the RME should be shown. In plots of risk (HQ or HI), the range from the LOAEL-based risk to the CTE receptor (CTE_L) to the NOAEL-based risk to the RME receptor (RME_N) should be shown using a format similar to that below:

CTE < ----------> RME	(use for plots of exposure)
CTE_L <----------> RME_N	(use for plots of HQ and HI)

Furthermore, all baseline risk assessments should be presented in a tabular format recommended in EPA *RAGS Part D* guidance (U.S. EPA, 1998b), which facilitates the review and interpretation of results.

Characterizing Variability in Exposure Dose

Figure 5-9 provides an example distribution of inter-individual variability in doses (mg/kg-day) to a population of receptors (e.g., some species of wildlife) exposed at a site. The top panel is a PDF, and the bottom panel is the same distribution plotted as a cumulative density function (CDF). The variability in dose may be a result of either a) variations between individuals in intake (different ingestion rates, different fraction of time spent in the contaminated area, etc.), and/or b) variability in concentrations at different home range locations within the site.

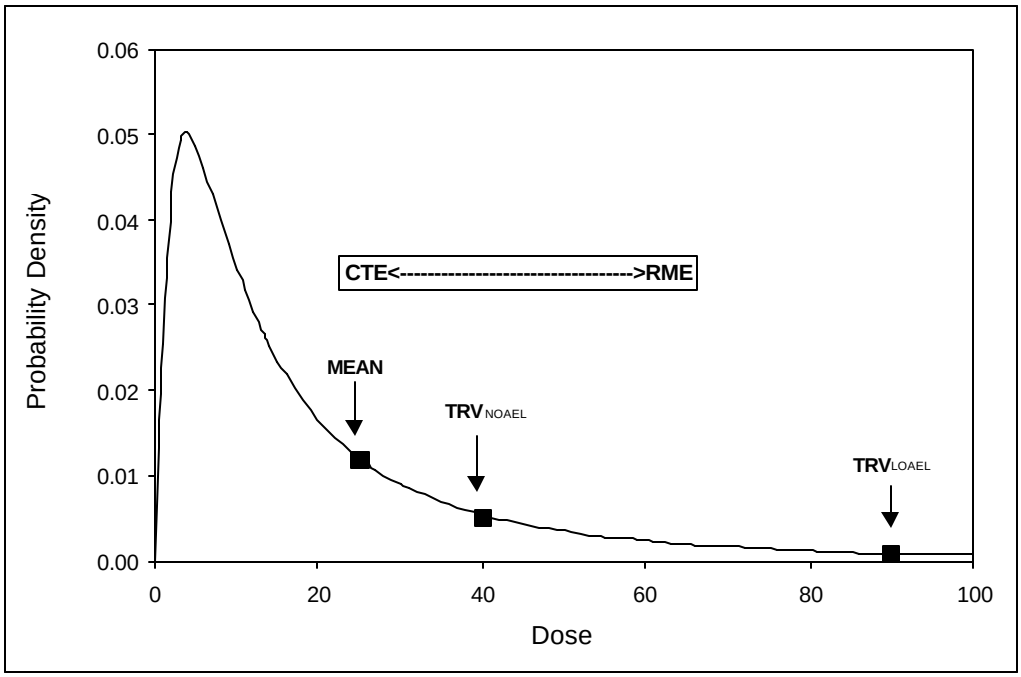
The information in graphs of this type may be used to characterize the range of hazards to different individuals in the receptor population by superimposing available information on the dose-response curve for the receptor and endpoint of concern. In most cases, this will take the form of one or more point estimates of dose-based TRVs. For example, the point estimates might be the estimated NOAEL and LOAEL for the receptor and endpoint of concern. This presentation allows quantification of the fraction of the population that is expected to receive doses above the NOAEL and above the LOAEL. For example, in this case the PRA of variability indicates that 84% of the population is exposed to doses below the NOAEL, 11% of the population is exposed to doses between the NOAEL and the LOAEL, and 5% is exposed to doses that exceed the LOAEL. A more advanced form of this analysis would be to superimpose the full dose-response curve used to select the NOAEL and the LOAEL, rather than just two point estimates derived from the curve. Figure 5-10 illustrates this approach.

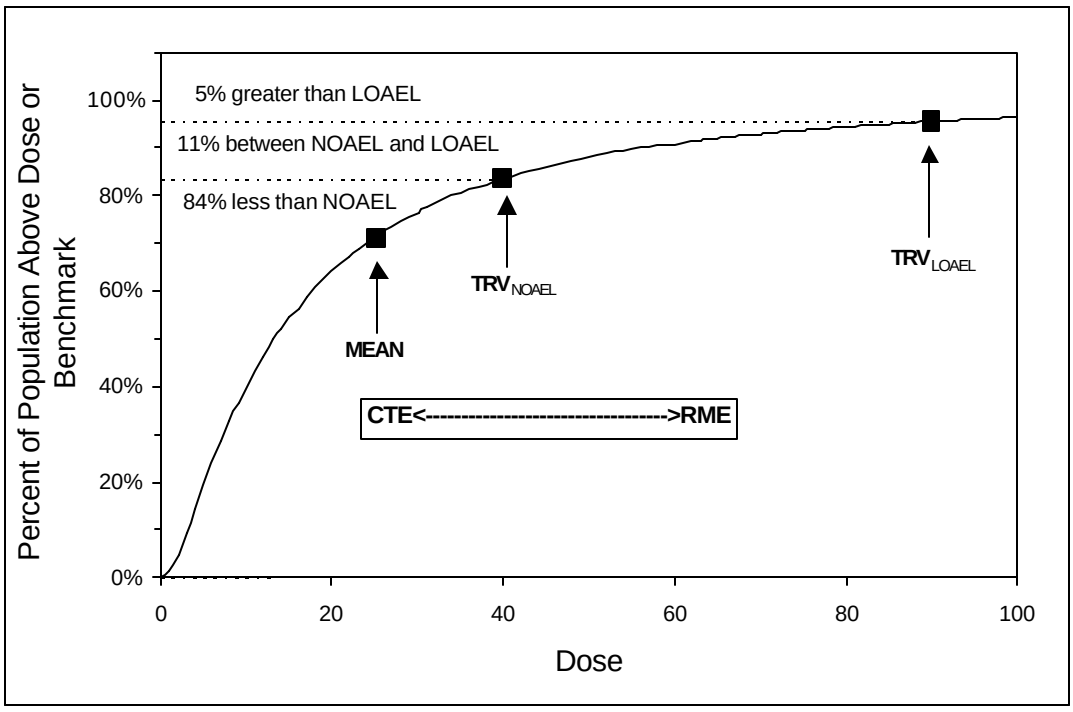
Characterizing Variability in Exposure Concentration

In some cases, TRVs that are used to evaluate risk are based on the concentration in a medium (e.g., soil, sediment, water) rather than the dose resulting from exposure to the medium. For example, this is most often the case for aquatic receptors (fish, benthic macroinvertebrates), terrestrial plants, and some terrestrial receptors (e.g., earthworms). Similar to the approach described above, the distribution of concentrations in the medium is plotted as a PDF or CDF, and one or more TRV's (expressed in units of concentration) are superimposed. This allows an evaluation of the fraction of all measurements that

exceed a level of concern. Note that a distribution for variability in concentration can generally be characterized using either a theoretical or an empirical distribution function (see Chapter 3).

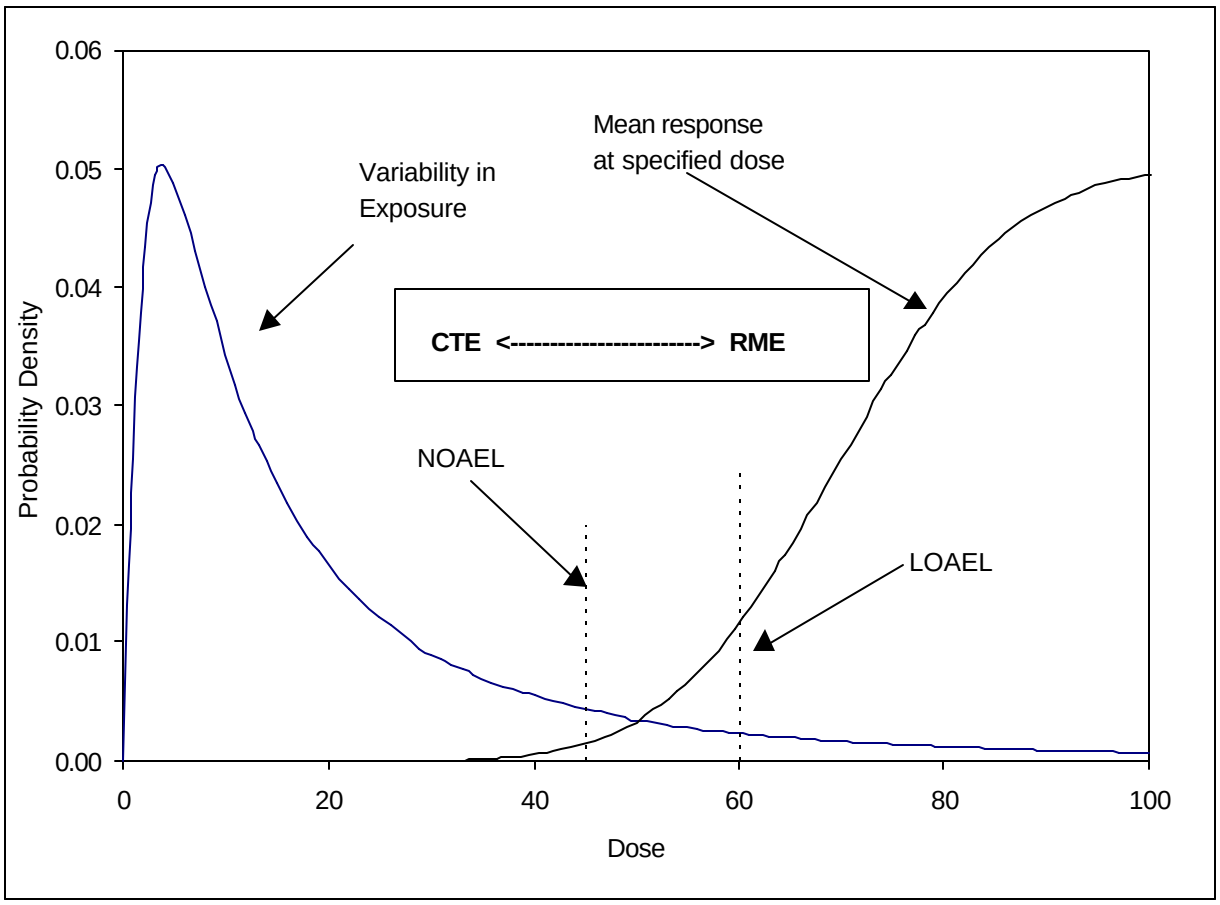
Figure 5-9. Example output of a PRA simulation of variability in dose among different individuals in a population. The upper panel is the PDF and the lower panel is the CDF. Point estimates of two toxicity reference values (the NOAEL and the LOAEL) are superimposed on the curves to indicate the fraction of the population that is likely to be exposed at levels of potential concern (greater than the NOAEL) or of probable concern (greater than the LOAEL).





1

Figure 5-10. Example output of a PRA simulation of variability in dose among different individuals in a population, with the full dose response curve (rather than just the NOAEL and the LOAEL) for some specified endpoint superimposed.



When the distribution of concentration values represents spatial variability (e.g., concentrations in soil at different locations within the site), the fraction of the distribution above some specified concentration-based TRV is an estimate of the areal extent of the site that exceeds the TRV. This approach is most useful when the receptors of concern are not mobile or have small home ranges . For example, this approach could be used to estimate the fraction of a site where soil concentrations may exceed levels that are phytotoxic to terrestrial plants or which cause adverse effects in soil invertebrates. Of course, this same information can be obtained by plotting all of the measured values on a map and using appropriate software techniques to estimate iso-concentration lines that are above or below some specified TRV.

In some cases, the distribution of concentration values may represent temporal variability in concentration in a medium at a location (e.g., the concentration in surface water at some particular sampling location in a stream or lake). In this case, the distribution of values may be used to characterize both the frequency and magnitude of temporal fluctuations which exceed some specified TRV. For example, this is often the approach taken when the TRV is an acute or chronic Ambient Water Quality Criterion (AWQC). The upper panel of Figure 5-11 illustrates this approach. In this example, the concentration of a contaminant in grab samples from a stream exceed the acute AWQC about 17% of the time. This type of analysis is mainly useful for organisms that are exposed at or near the sampling station, and is not immediately applicable to organisms that are mobile and may be exposed at widely separated locations where concentration values may be different.

A variation on this approach is shown in the bottom panel in Figure 5-11. In this case, the distribution of concentration values is not compared with one or more point estimates of a TRV in a specific receptor, but with a distribution of TRVs for a range of different receptors that may be exposed to that medium For example, consider a number of different species of aquatic organism (benthic invertebrates, fish, amphibians, etc.) that all reside in a particular location (e.g., a stream or lake) that has time-variable concentrations of a contaminant in water. In this case, the results indicate that concentration values would occasionally reach levels that could affect some of the most sensitive species of receptors, but that most species would not be affected.

Characterizing Variability in Hazard Quotient

Figure 5-12 shows hypothetical distributions for Hazard Quotient (HQ) values calculated for a receptor population at a site. This type of presentation allows an assessment of the fraction of the population that is expected to have an HQ value within a specified range (e.g., 0–1, 1–2, 2–4, >4, etc.). As discussed below, interpretation of the distribution of HQ values depends on the toxicological basis of the TRV, but values below 1.0 are generally associated with little or no risk, while values above one are associated with the potential occurrence of adverse effects in individuals and/or populations.

Figure 5-11. Example of methods used to evaluate ecological risk when the toxicity reference values (TRVs) are in units of concentration rather than dose. This is usually the situation for assessment of exposures of aquatic receptors. (Top) Hypothetical distribution of concentration values of a contaminant in water at a particular sampling location. The distribution indicates 17% of all measurements exceed the TRV (the Ambient Water Quality Criterion). (Bottom) The PDF on the left is the distribution of concentrations measured at different times at a sampling location. The PDF on the right shows the range of toxicity reference values for different aquatic species exposed to this medium. In this example, sensitive species (those with TRVs in the range of 20-50 mg/L) are likely to be significantly affected, while species with TRVs above 60-80 mg/L are likely to have relatively low risk of effect.

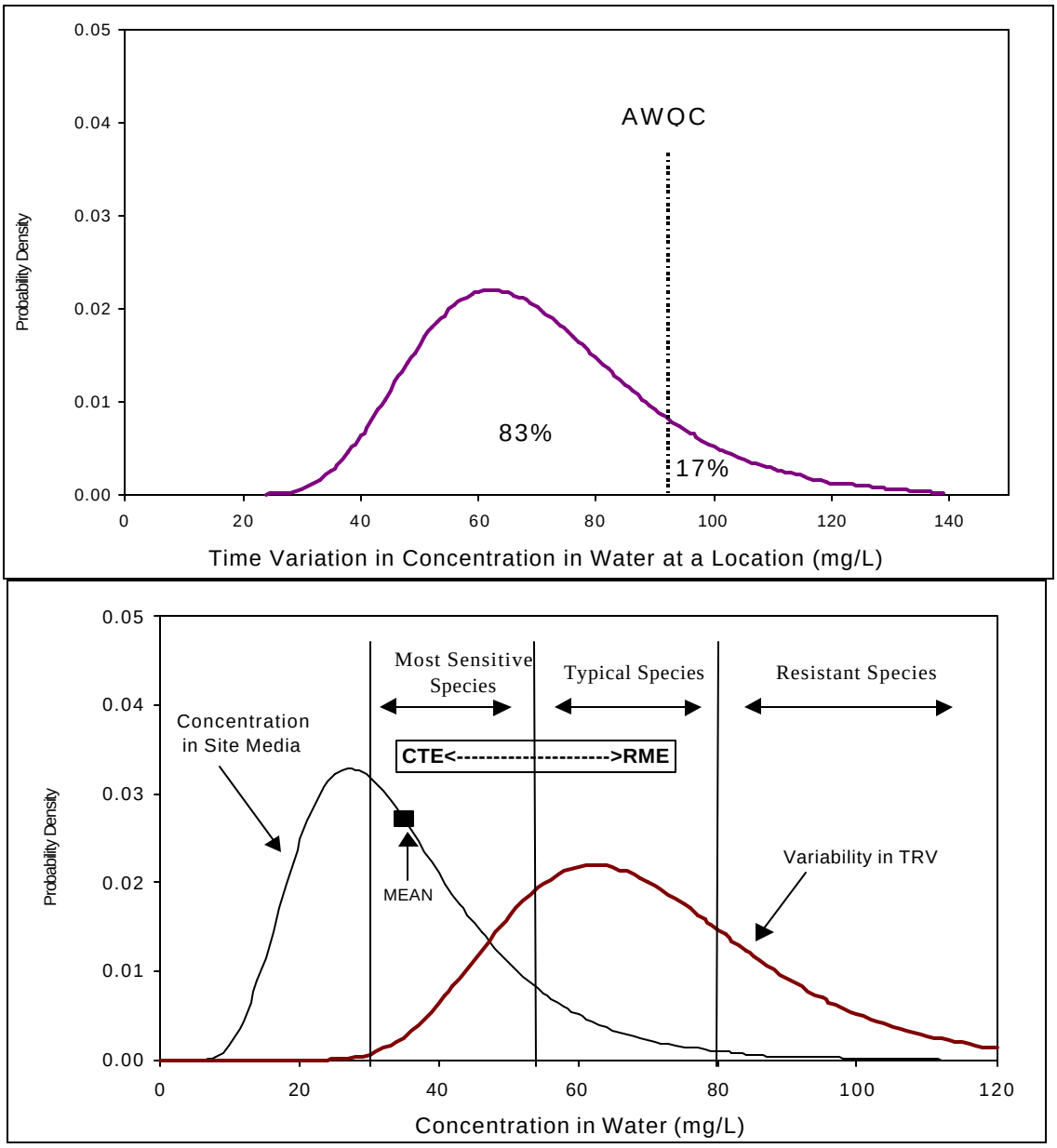
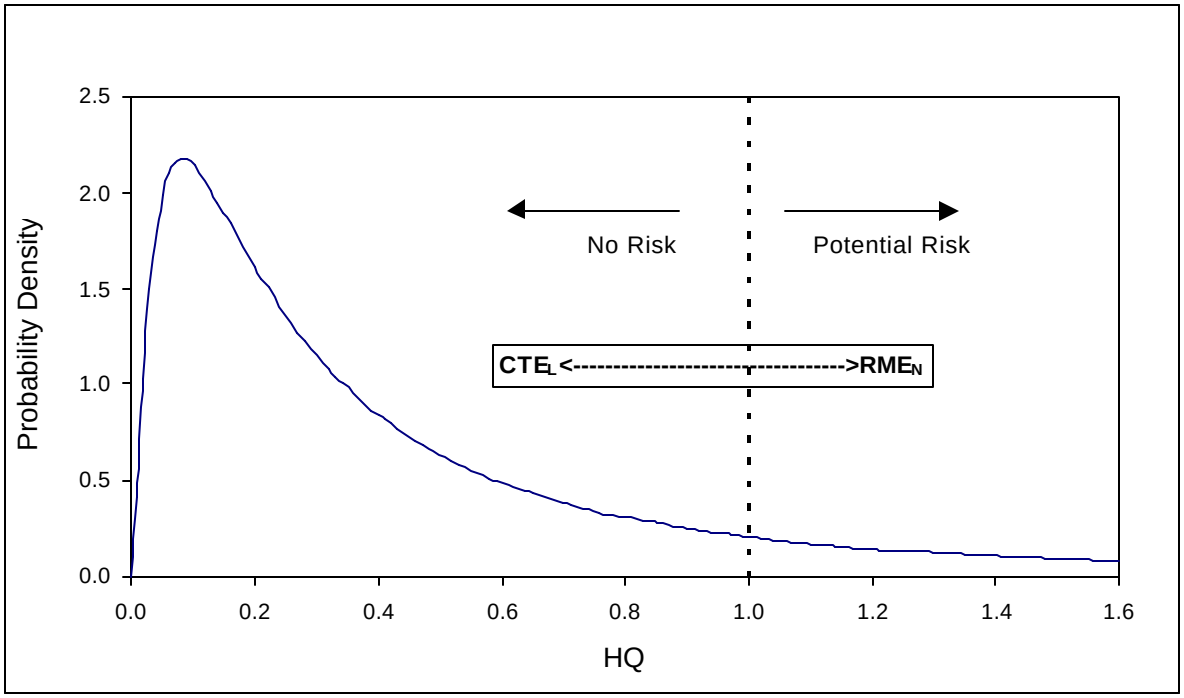
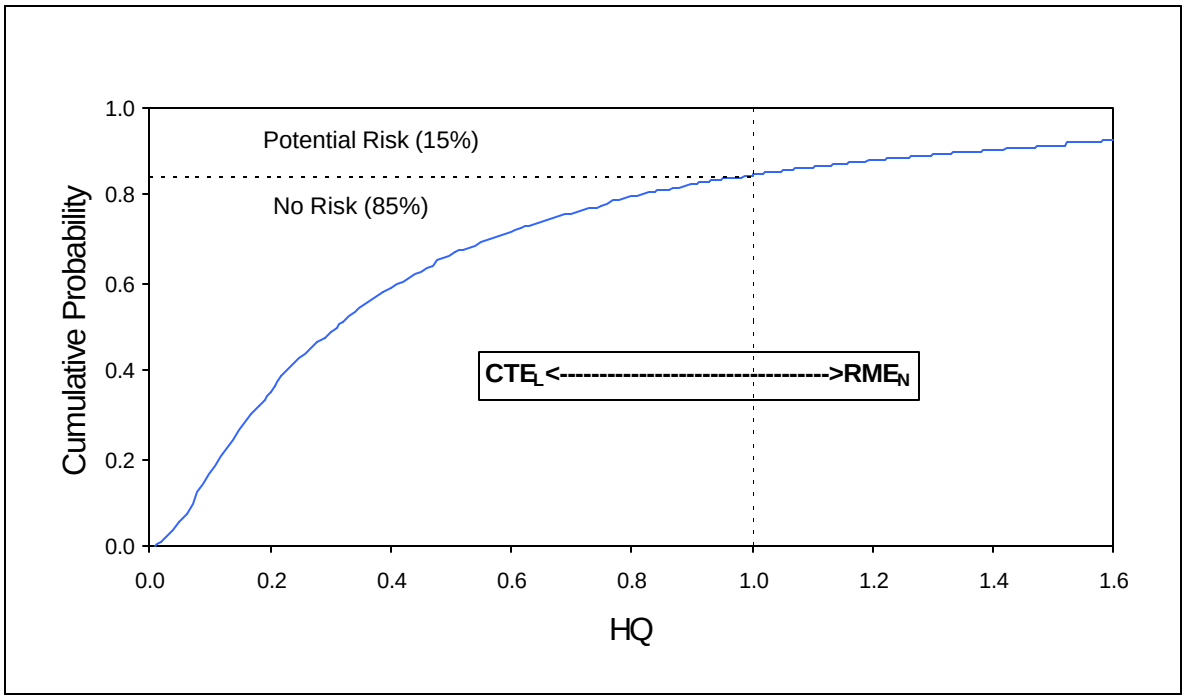


Figure 5-12. Example output of a PRA simulation of HQ across different individuals in a population. The upper panel is the PDF and the lower panel is the CDF. Using an HQ of 1.0 as the frame of reference, if can be seen that about 85% of the population is exposed at levels that are not of concern, and that about 15% of the population has exposures that may be of concern.





Characterizing Variability in Response

If inter-individual variability in the magnitude or severity of a response at a dose (concentration) and across doses (concentrations) can be modeled with PDFs (see Section 5.1.3 and Figures 5-6 or Figure 5-7), then Monte Carlo simulation techniques can used to quantify the distribution of responses rather than doses or HQs. An example output of such a simulation is shown in Figure 5-13. This type of distribution is conceptually similar to a distribution of doses or HQ values, except that more information is conveyed. For example, based on the nature of the endpoint, the magnitude of the response may be categorized as either minimal, moderate or severe, and the distribution reveals what fraction of the population is expected to fall within each category.

5.3.2 INTERPRETING DISTRIBUTIONS OF VARIABILITY

In contrast to the case for human health risk assessments (where default risk-based decision guidelines are well established), there are no established default decision guidelines for identifying when risks to ecological receptors are of concern. However, if risks to nearly all members of the population (including the RME individual) are below an appropriate NOAEL-based TRV, then it is likely that risks are within an acceptable range, both for individuals and the population. Conversely, if risks to average members of a population exceed an appropriate LOAEL-based TRV, then it is likely that risks are not acceptable, either to individuals or the population. For cases that fall between these bounding conditions, the level of effect that is considered acceptable generally should be defined by the risk assessor and the risk manager on a site-specific and receptor-specific basis by considering the following:

1. The risk management goal

If the risk management goal for the receptor being evaluated is protection of individuals (e.g., when the receptor is a threatened or endangered species), then risks should generally be considered acceptable if none or only a small percentile of the distribution exceeds a level of concern. In the case when the risk management objective is population sustainability, the decision is more complex, since adverse effects on some members of the population may not lead to an unacceptable impact on the population.

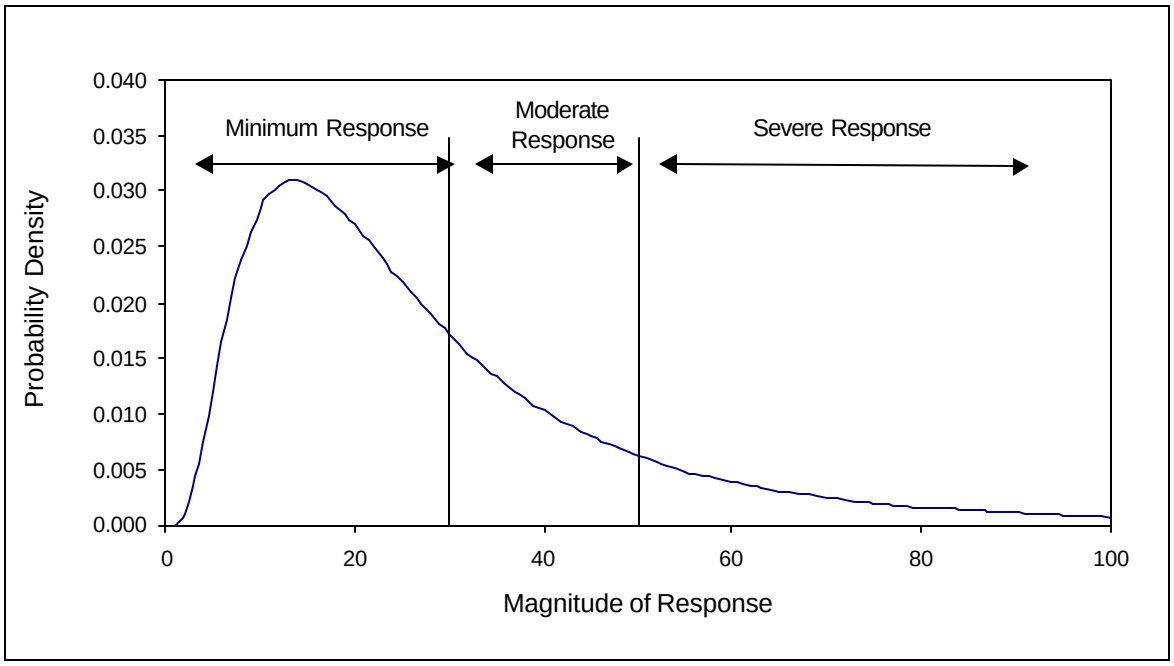
2. The toxicological basis of the TRV

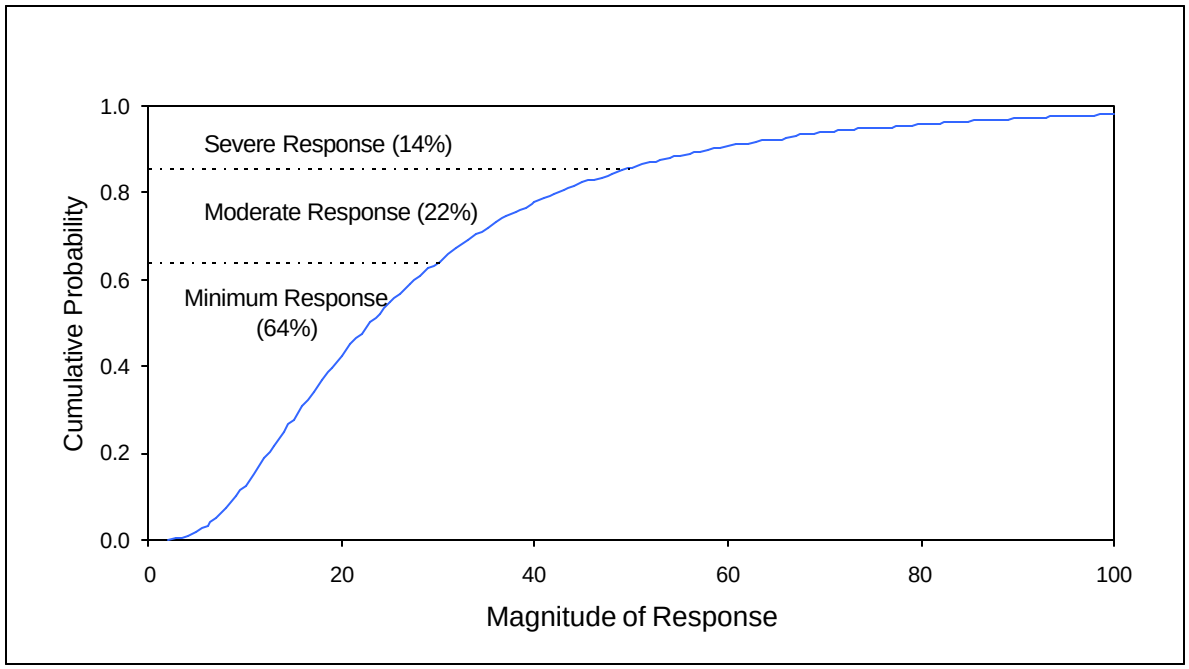
In order to understand the biological significance of a distribution of variability in dose or HQ, it is important to understand the nature of the TRV being used to evaluate the distribution. This includes the nature of the endpoint, its relevance to the assessment endpoint, and the shape (steepness) of the dose-response curve. For example, an HQ of 1.5 based on an EC20 for reduction in reproductive success would likely be interpreted as more significant than an HQ of 1.5 based on the EC20 for an increase in liver weight. Likewise, an HQ of 1.5 based on an EC0 for acute lethality would generally be more significant if the dose-response curve for lethality was steep than if it was shallow.

3. The characteristics of the receptor

Ultimately, the question that should be addressed is whether an effect of degree "x" occurring in "y" percent of the population is biologically and ecologically significant. This, in turn, depends on the attributes of the receptor being evaluated. For example, a reduction of 10% in the reproductive success of a fecund and common species (e.g., the field mouse) might not lead to a significant reduction in population number, while the same effect could be of concern in a species with lower fecundity and/or lower population density (e.g., the moose). Thus, the interpretation of an analysis of variability in exposure and/or effect would often benefit from the input of a trained population biologist with expertise in the receptor of concern.

Figure 5-13. Example output of inter-individual variability in the magnitude of the toxicological response to a chemical stressor in different members of a population. The upper panel is the PDF and the lower panel is the CDF. As seen, about 64% of the population is expected to have no or minimal response, about 22% is expected to have a moderate response, and about 14% is expected to have a severe response. This type of risk characterization is helpful in estimating whether population-level effects are likely to occur.





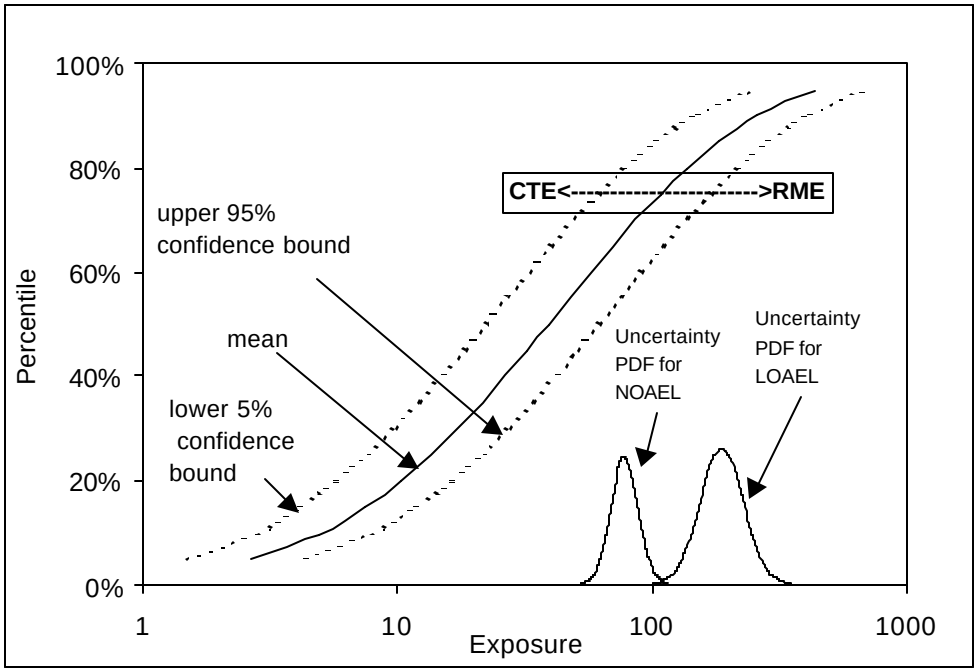
5.3.3 PRESENTING AND INTERPRETING DESCRIPTIONS OF UNCERTAINTY

There are several different techniques for presenting information on uncertainty. In the most common case (when a 2-D Monte Carlo simulation has been run to characterize both variability and uncertainty in some parameter such as exposure or risk), the results may be conveniently displayed in tabular format, similar to that illustrated in Table 5-1. When graphic formats are used, it is generally easiest to use CDFs rather than PDFs to display the results, similar to those shown in Figure 5-14. Such graphic formats of exposure distributions also allow superimposition of relevant TRVs such as the point estimates and/ or the uncertainty distributions of the NOAEL and the LOAEL for a particular endpoint (Figure 5-14, upper panel). If the TRVs are treated as uncertain, then it is usually best to plot the variability distribution and confidence bounds of the HQ rather than attempting to display the variability and uncertainty of both the exposure and the TRV in a single graph (Figure 5-14, lower panel).

In the case when there is a clear decision guidance for deciding when risk is acceptable or unacceptable, then a display of the variability around that decision rule may be helpful. For example, assume that a population biologist at a site has determined that if less than 15% of a population of field mice are exposed to doses that exceed the EC20 for reduction in pups per litter, then population level effects are not expected at that site. Thus, the percent of the population exposed above the EC20 is the "critical statistic" that is needed for decision making. A Monte Carlo analysis of variability in dose indicates that 11% of the population of field mice are exposed at levels above the EC20. Based on this, it is concluded that exposure levels at the site are not likely to cause population-level effects and so would be acceptable. However, the estimate that only 11% of the population exceeds the EC20 level is not certain, and there is a chance that the true percentage of the population above the EC20 actually exceeds the 15% level. Figure 5-15 illustrates a graphical format that is useful for displaying information on the uncertainty around a critical risk statistic. Two hypothetical distributions are shown, both with the same best estimate value (11%). In one case, the uncertainty bound is relatively narrow, and there is only a small chance that the actual fraction of the population above the EC20 is higher than 15%. In the second case, the uncertainty bounds are quite wide, and there is a large chance that at least 15% of the population could be exposed above the EC20.

There are no default rules for how to utilize information on uncertainty bounds when making risk management decisions. In most cases, the best estimate of the critical statistic (i.e., the CTE of the uncertainty distribution around the critical statistic) should be given greatest credence, since it is more likely to be correct than some other estimate. However, values either higher or lower in the uncertainty range may be employed for decision-making (along with other relevant factors). For example, the wider the uncertainty range (and the greater the probability that the true risk is substantially higher than the best estimate), the more reasonable it is for a risk manager to base a decision on the upper portion of the uncertainty distribution. Likewise, the more serious the consequences of making an error, the more reasonable it is to make decisions based on the upper end of the uncertainty distribution. This is especially true if the dose-response curve for the effect of concern is steep (i.e., a small misjudgment in the level of exposure could be associated with a large difference in the effect on individuals and/or the

Figure 5-14. (Top) Example plot of the variability and uncertainty in exposure. The CDF format is used to help keep the figure simple. The scale may be either linear or log (as in this example). TRVs may be superimposed, either as point estimates, or as variability or uncertainty distributions (as shown in the example). (Bottom) An example plot of variability and uncertainty in HQ or HI.



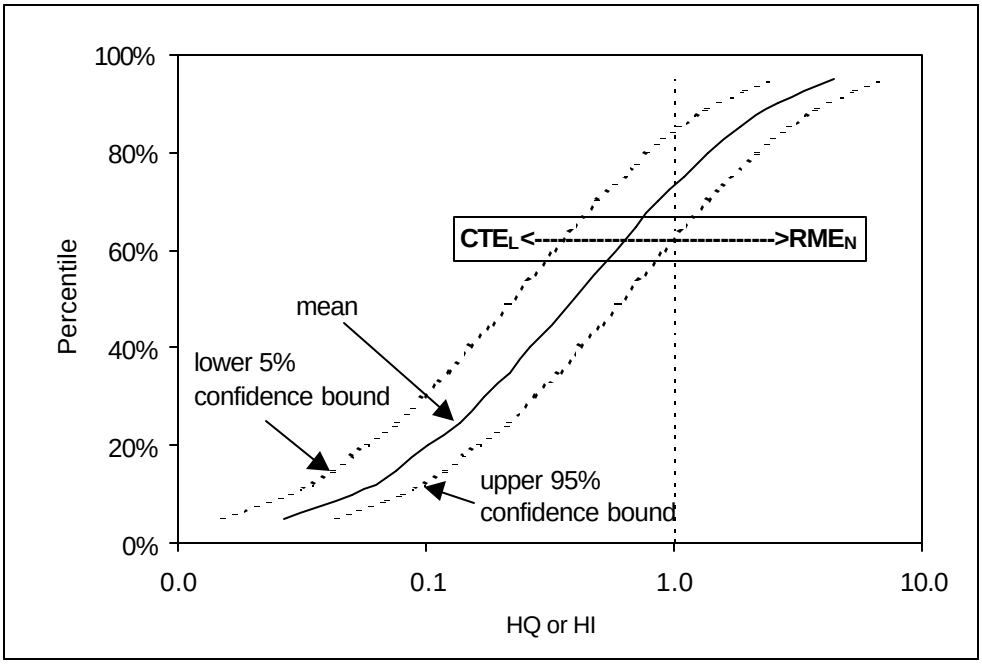
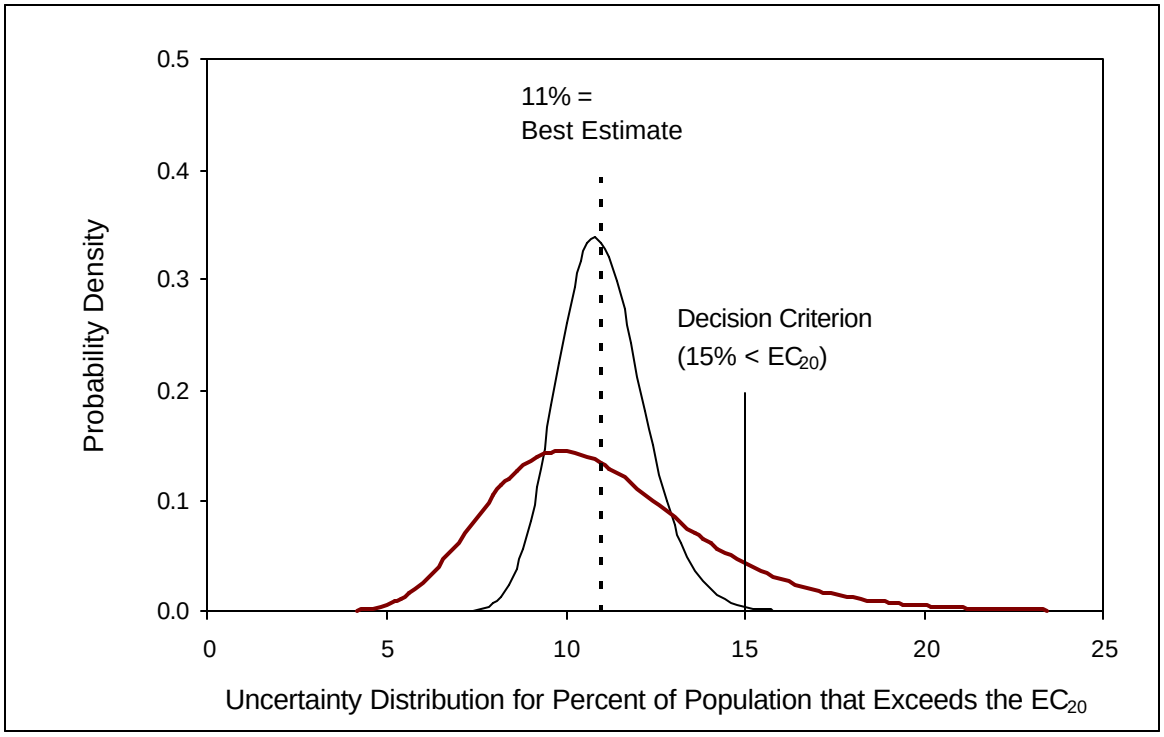


Table 5-1. Example of tabular format for displaying the results of a 2-D Monte Carlo simulation. These data can also be presented graphically, using a format similar to that shown in Figure 5-14.

Variability Percentile	Uncertainty Percentile					
	0.05	0.25	0.50	0.75	0.95	mean
0.05	1.5	2.0	2.5	3.2	4.4	2.7
0.10	2.8	3.8	4.7	5.9	8.0	5.0
0.15	4.2	5.7	7.1	8.8	12.0	7.5
0.20	5.8	7.9	9.8	12.2	16.6	10.3
0.25	7.7	10.4	12.9	15.9	21.8	13.6
0.30	9.8	13.2	16.4	20.3	27.6	17.3
0.35	12.3	16.5	20.5	25.3	34.4	21.5
0.40	15.1	20.3	25.2	31.2	42.2	26.5
0.45	18.5	24.8	30.8	38.0	51.6	32.3
0.50	22.4	30.1	37.3	46.1	62.5	39.3
0.55	27.2	36.6	45.3	55.9	75.7	47.6
0.60	33.1	44.4	55.1	67.6	91.4	57.8
0.65	40.3	54.2	67.2	82.8	111.5	70.4
0.70	49.6	66.9	82.8	101.9	137.1	86.7
0.75	62.0	83.2	103.3	127.7	171.3	108.3
0.80	79.1	106.5	132.0	162.7	219.2	138.4
0.85	104.9	141.5	175.2	215.8	290.0	183.6
0.90	149.6	200.9	249.4	307.6	414.3	261.4
0.95	250.7	336.5	416.9	516.9	695.3	438.3
Mean	65.3	87.2	107.8	133.4	179.0	113.0

Figure 5-15. Example format for displaying the uncertainty range around the "critical statistic" that will be used for decision making. In this case, the critical statistic is the percent of the population whose exposures exceed the EC20 for reproductive effects. Two example uncertainty distributions are shown, both with the same best estimate (11%). In one case, the uncertainty range is relatively narrow, and in the other case the uncertainty range is wide.



population). In some cases, the uncertainty distribution may be so wide (spanning a range from far below to far above a level of concern) that a risk manager may decide that there is no appropriate basis to make a decision. This result would support a recommendation to collect additional data to narrow the uncertainty range.

5.3.4 COMBINING PRA RESULTS WITH OTHER LINES OF EVIDENCE

By definition, risk characterization for an ecological risk assessment includes two phases: risk estimation and risk description (U.S. EPA, 1997a, 1998a). Risk description is defined as the interpretation and discussion of the available information about quantitative risks related to the assessment endpoints, including a discussion of the lines of evidence supporting or refuting the risk estimate(s) and an interpretation of the significance of the risks related to the assessment endpoints. This may include, for example, site-specific measurements of chemical exposure or toxicity (e.g., soil, water and sediment toxicity testing and measurements of benthic invertebrate community structure and function), as well as site-specific surveys of receptor diversity and/or abundance. The risk assessor establishes a relationship between the assessment endpoints and measures of effect and associated lines of evidence in quantifiable and clearly described terms. PRA results should help to establish these relationships and provide part of the basis for clear, transparent, reasonable, and consistent risk characterization recommended by EPA policy (U.S. EPA, 1995a). The risk assessment process is iterative, and depends upon how much confidence risk managers have in making informed decisions based upon the uncertainties in the risk assessment. If uncertainties are too great, then additional iterations of data collection can be pursued to fill data-gaps and reduce uncertainties.

5.3.5 COMMUNICATING THE RESULTS OF PRA ANALYSES

Risk communication (see Chapter 8) is generally the responsibility of the risk manager but may be shared with risk assessors (U.S. EPA, 1999). The goals of risk communication generally include:

- A clear description of the source(s) and cause(s) of risks
- The ecological relevance of the assessment endpoint(s) (e.g., earthworms may not be perceived as important by the public, but may play a critical role in the food web (U.S. EPA, 1999))
- The potential adversity of the risk (e.g., nature and intensity, spatial and temporal scale and recovery potential)
- The degree of confidence in the risk estimation and risk description
- The rationale for the risk management decision
- The remedial alternatives for reducing risk

The PRA results and the associated graphs illustrating the confidence about the risk estimation can be useful for risk communication. Graphs help to describe confidence intervals and relative contributions of

exposure to risk. These tools can prove equally useful for the evaluation of remedial alternatives if the risk estimates can be linked to associated costs of remediation in order to provide improved cost/benefit information to the risk manager.

5.4 GENERAL GUIDELINES FOR SUBMISSION OF A PROBABILISTIC ECOLOGICAL RISK ASSESSMENT

EPA has issued general non-program specific guidance on the use of Monte Carlo Analysis in risk assessment as *Guiding Principles for Monte Carlo Analysis ("Principles")*(U.S. EPA, 1997b) (see Chapter 1). As part of this guidance, EPA provides sixteen guiding principles for Monte Carlo Analysis. These principles describe the key elements of a successful MCA and each should generally be considered equally important to the success of PRA for ecological risk assessments.

Dialogue among Stakeholders

As discussed in Section 5.2, the risk assessors and risk managers should begin a dialogue concerning the potential application of the PRA to the ecological risk assessment. An initial scoping meeting can be used to discuss the goals of the PRA and the potential value of the analyses. Similar meetings are expected to occur when a proposal is made to move the PRA to a higher level of complexity.

The stakeholders who would participate in discussions may include the members of the public, representatives from state or county environmental agencies, tribal government representatives, natural resource trustees, and potentially responsible parties (PRPs) and their representatives (U.S. EPA, 1999).

Preparation of the Work Plan

A work plan should generally be submitted by the contractor or PRP to the BTAG coordinator and/or regional ecotoxicologist for review and for approval by the risk manager. EPA strongly recommends that PRPs involve the Agency in the development of a workplan prior to commencing the analysis. The work plan for a PRA is discussed in detail in Chapter 6 and highlighted in the Exhibit 5-6.

The work plan for the PRA should be developed according to available guidance for workplans for point estimate ecological risk assessment (U.S. EPA, 1992b, 1997a) and generally should consider three elements: 1) the sixteen guiding principles of MCA (U.S. EPA, 1997b); 2) the eight guidelines for PRA report submission (U.S. EPA, 1997b; see Chapter 1, Section 1.7 and 3) the tiered approach to ecological risk assessment (U.S. EPA, 1997a; 1999).

The PRA Report

The *Principles* guidelines explicitly identify eight conditions for submission of a successful PRA. These conditions (detailed in Chapter 4) should be considered when designing and reporting results of a PRA for Superfund. A checklist for reviewing a PRA is discussed in detail in Chapter 6 (see Section 6.2) and highlighted in Exhibit 5-7.

Internal and External Review of the PRA Workplan and Report

At the discretion of the EPA risk assessor or risk manager, the PRA work plan and report may be submitted for additional EPA internal review and/or an external review process. EPA has new national and regional guidances for conducting peer reviews (U.S. EPA, 1998b). The external peer review may be used in cases where the issues are complex or contentious and the opinions of outside expert peer review

EXHIBIT 5-6

WORKPLAN FOR A PRA

- A statement of the assessment and measurement endpoints (or "measures of effect", U.S. EPA, 1998a) and receptor's of ecological concern
- A depiction of the site conceptual model and food chain exposure models
- Summary of the point estimate risk assessment and rationale for the PRA
- Description of the methods and models to be used for the PRA
- Initial sensitivity analyses
- Preliminary Monte Carlo Simulation
- Refined sensitivity analyses and discussion of influential and uncertain variables
- Proposal for obtaining and using distributions
- Methods for deriving the exposure term(s)
- Methods for deriving the toxicity term(s)

EXHIBIT 5-7

CHECKLIST FOR INCLUDING A PRA AS PART OF THE ERA (SEE CHAPTER 6)

- In general, risk assessments should include point estimates prepared according to current Superfund national (preemptive) and regional (clarifying) guidance.

- A workplan should be submitted for review and approval by the appropriate EPA regional office risk manager (RPM) and/or Biological Technical Assistance Group (BTAG) coordinator prior to submission of the PRA.

- A tiered approach should be used to determine the level of complexity appropriate for the ecological risk assessment. The decision to ascend to a higher level of complexity generally should be made with the risk manager, regional risk assessor and other stakeholders.

- The eight conditions presented in the EPA policy on PRA (U.S. EPA, 1997b; see Chapter 1, Section 1.7) generally should be addressed by each PRA submitted to the Agency.

- Information in the PRA generally should possess sufficient detail that a reviewer can recreate

1

EXHIBIT 5-8

EXAMPLE OF INPUTS FOR A GENERAL RISK MODEL FOR A REPRESENTATIVE WILDLIFE SPECIES

HQ = Dose ÷ TRV

Dose (mg/kg-day) = Range from central tendency (CTE) to upper end (RME)

Dose (mg/kg-day) = C @ (IR / BW) @ BA @ AUF

where:

C (mg/kg)	=	Range of concentration values (CTE to upper end) in an environmental medium , or PDF fit to measured values at the site
Intake (kg/day)	=	Range of point estimates of intake of environmental medium (CTE to upper end), or PDF fit to measured or literature values
BW (kg)	=	Range of point estimates of body weight (CTE to upper end), or PDF fit to measured or literature values
BA	=	Range of point estimates of bioavailability (CTE to upper end), or PDF fit to measured or literature values
AUF	=	Area use factor (fraction of time spent in exposure unit), either as a point estimate or as a PDF
TRV (mg/kg-day)	=	Toxicity Reference Value; either a point estimate of NOAEL, LOAEL, BMD, etc., or a PDF which characterizes variability or uncertainty in the TRV

can improve the PRA. The external peer reviewers should possess no bias or agenda concerning the process or methods of the PRA and they should have no stake in the outcome of the risk assessment. When reviewing a PRA report, the risk assessor should ensure that the tiered approach was followed and that the risk assessment conforms to the details of the workplan. Both the checklist above and EPA's guiding principles (U.S. EPA, 1997b) may be consulted.

Changes in Scope after Completion and Approval of Work Plan

The PRA generally should be completed as an iterative process moving from a simple level of complexity (receptors, number of variables) to a more complex level of effort (additional receptors and number of variables) as is indicated for an adequate site risk assessment. As new information is developed, it needs to be considered with the future possibility of moving to a higher level of complexity. The decision to use a more complex approach for quantifying variability and uncertainty generally should be accompanied by a revised workplan.

Selection and Fitting of Distributions

Identification of the sources of information data for the input distributions is one of the most important parts of the work plan process for the PRA, since erroneous inputs can create equally or more erroneous outputs for risk assessment. Chapter 3 provides guidance on the selection of data and the fitting of distributions to the data. Much of this guidance is fully applicable to the use of PRA in ecological risk assessments. Special considerations regarding the selection and fitting of distributions for ecological evaluations include the following:

- It is possible to evaluate both inter-individual variability and/or uncertainty in toxicity factors (TRVs).

- The sources of information for input distributions are different. For example, the most convenient source of exposure factors for wildlife is the *Wildlife Exposure Factors Handbook* (U.S. EPA, 1993a). In some cases, exposure data are very limited, and this may pose a difficulty to deriving reliable PDFs.

- It is not always as important to model inter-individual variability in exposure and risk in an ecological risk assessment as in a human health risk assessment. In some cases, when risk decisions can be based on a description of risk to average (rather than RME) individuals in a population, it may be more helpful to model uncertainty in the estimate average (CTE) exposure or risk using one-dimensional methods.

In general, selecting and fitting PDFs for variability in exposure factors and toxicity benchmarks is one of the more challenging aspects of proceeding with a good PRA. One can use default, site-specific or a combination of these distributions as well as point estimates in PRAs. The EPA *Wildlife Exposure Factors Handbook* (U.S. EPA, 1993a) is generally the preferred source of inputs for both point estimate and PRA based ERAs for terrestrial wildlife. Site-specific data can be used as inputs when available and to the degree usable. Exhibit 5-8 identifies some common input variables often needed for ERAs, based on a generic HQ (risk) equation utilizing a dose-based TRV. The inputs may be either

point estimates or distributions with shapes and ranges that are supported by literature or by site-specific exposure and/or toxicity data. Common shapes are normal, lognormal, triangular, and uniform; minimums and maximums generally should represent plausible extremes based on available site-specific data and judgment. Software programs can help fit distributions to data with increasing statistical power. Deviations from any default values or ranges specified by guidance generally should be supported with defensible site-specific data.

Knowledge of the relative homogeneity or heterogeneity of the contaminant concentrations within an Exposure Unit Area (often related to home ranges) is essential in deriving useful PDFs for wildlife exposure. The utility of analytical data (soils, sediment, surface water or tissues) generally should be evaluated in terms of the representativeness of samples in characterizing chronic exposures. The randomness and density of sample locations, media characteristics related to intake, integrity of samples, and proper analytical methods may be important considerations. Sensitivity analyses should be performed on the inputs to the risk model to determine which factors have the greatest influence on the output distributions. Data-gaps can generally be prioritized for sampling and analyses in an attempt to reduce uncertainties in the risk assessment.

REFERENCES FOR CHAPTER 5

National Academy of Sciences (NAS). 1983. *Risk Assessment in the Federal Government: Managing the Process*. Committee on the Institutional Means for Assessment of Risks to Public Health, National Research Council. 1983.

National Academy of Sciences (NAS). 1994. *Science and Judgment in Risk Assessment*. Committee on Risk Assessment of Hazardous Air Pollutants, National Research Council. 672 pages.

U.S. EPA. 1989a. *Risk Assessment Guidance for Superfund Volume II Environmental Evaluation Manual. Interim Final. Office of Emergency and Remedial Response*. Washington, D.C. EPA/540/1-89/001. March 1989.

U.S. EPA. 1989b. *Risk Assessment Guidance for Superfund.* (RAGS): *Volume II. Environmental Evaluation Manual.* Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/001.

U.S. EPA. 1990. *National Oil and Hazardous Substances Pollution Contingency Plan*, Final Rule. 40 CFR 300.430.

U.S. EPA. 1991. *Eco Update*. Intermittent Bulletin Series. Office of Emergency and Remedial Response. 1991 to present.

U.S. EPA. 1992a. *Framework for Ecological Risk Assessment.* EPA Risk Assessment Forum. EPA/630/R-92/001. February 1992.

U.S. EPA. 1992b**.** Policy Memorandum: *Guidance on Risk Characterization for Risk Managers and Risk Assessors* from F. Henry Habicht, Deputy Administrator, February 26, 1992. OSWER Directive 9285.7-17.

U.S. EPA. 1992c. *Guidance for Data Usability in Risk Assessment (Part A).* Office of Emergency and Remedial Response, Publication 9285.7-09A, PB92-963356, April 1992.

U.S. EPA. 1993a. *Wildlife Exposure Factors Handbook* Volumes 1 and II. EPA/600/R-93/187a and b. Office of Research and Development. December.

U.S. EPA. 1993b. *Data Quality Objectives for Superfund: Interim Final Guidance.* Office of Emergency and Remedial Response, Publication 9255.9-01, Washington, DC. EPA/540-R-93-071. NTIS PB94-963203.

U.S. EPA. 1994. Memorandum*: Role of the Ecological Risk Assessment in the Baseline Risk Assessment*. From Elliott Laws, Assistant Administrator to Directors, Waste Management

Divisions and Directors, Environmental Services Division. OSWER Directive No. 9285.7-17. August 12, 1994.

U.S. EPA. 1995a. *EPA Risk Characterization Program*. Memorandum from the Administrator. March 21, 1995.

U.S. EPA. 1995b. The Use of the Benchmark Dose Approach in Health Risk Assessment. Office of Research and Development. Publication EPA/630/R-94/007. February 1995.

U.S. EPA. 1997a. *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments.* Interim Final. U.S. Environmental Protection Agency, Environmental Response Team (Edison, NJ). June 5, 1997.

U.S. EPA. 1997b. *Guiding Principles for Monte Carlo Analyses*. EPA/630/R-97-001.

U.S. EPA. 1998a. *Guidelines for Ecological Risk Assessment.* Risk Assessment Forum. U.S. Environmental Protection Agency, Washington DC. EPA/630/R-95/002F. April. May 14, 1998 Federal Register 63(93): 26846-26924.

U.S. EPA. 1998b. *Risk Assessment Guidance for Superfund: Volume 1 – Human Health Evaluation Manual (Part D, Standardized Planning, Reporting and Review of Superfund Risk Assessments)*. Interim. EPA 540-R-97-033. OSWER 9285.7-01D, PB97-963305, http://www.epa.gov. January 1998.

U.S. EPA. 1999. *Ecological Risk Assessment and Risk Management Principles for Superfund Sites*. OSWER 9285.7-28P. October 1999.

CHAPTER 6

WORKPLAN AND CHECKLIST FOR PRA

6.0 INTRODUCTION

This chapter outlines the basic concepts associated with developing a workplan prior to the initiation of a probabilistic risk assessment (PRA), and using a checklist when reviewing a PRA. Like the quality assurance project plan (QAPP), the workplan for PRA generally should document the combined decisions or positions of the RPM, risk assessor, and stakeholders involved in the risk assessment. There are many stakeholders in a risk assessment (see Chapter 1, Section 1.6), and it is important to involve and engage all stakeholders early in the decision-making process. These are important steps that will save time and effort.

6.1 WORKPLAN

In general, PRAs may be developed by EPA, EPA contractors, or a potentially responsible party. In each case, it is important to develop a workplan early in the risk assessment process. PRA's to be submitted by a contractor or potentially responsible party (PRP) should generally be submitted for EPA review before commencing the analysis. The workplan should describe the software to be used, the exposure routes and models, and input probability distributions and their basis (i.e., relevance to the site-specific contamination and pathways), including appropriate literature references. Examples of contents of the workplan are given in Exhibit 6-1. It is important that the risk assessor and risk manager discuss the scope of the probabilistic analysis and the potential impact on the Remedial Investigation/Feasibility Study.

- *Given the time and effort that can be expected to be invested in conducting a PRA, it is important*

EXHIBIT 6-1

CONTENTS OF THE WORKPLAN

1. Statement of the ecological assessment endpoints and/or human risk
2. Value added by conducting a PRA and proceeding to the subsequent tiers
3. Summary of the point estimate risk assessment and the rationale for the PRA
4. Discussion of adequacy of environmental sampling for PRA or moving to a successive tier (e.g., data quality issues)
5. Description of the methods and models to be used (e.g., model and parameter selection criteria)
6. Proposal for obtaining and basis for using exposure factor distributions
7. Methods for deriving the concentration term
8. Initial sensitivity analysis
9. Preliminary Monte Carlo simulation
10. Refined sensitivity analysis and discussion of influential and uncertain variables
11. Software (i.e., date and version of product, random number generator)
12. Proposed schedule, discussion points, etc.

that a workplan undergo review and approval by EPA, prior to proceeding with the assessment.

EPA generally will not accept probabilistic analysis where a workplan for the analysis has not been initially submitted to the Agency and approved by the Regional risk assessor and RPM. Exceptions to this process may be considered on a case-by-case basis. As the assessment moves to a higher tier, the workplan should be revised or a new workplan developed. The assessment should generally be restricted to the chemicals and pathways of concern that contribute the greatest risk.

Conducting a PRA is an iterative process. In general, as new information becomes available, it should be used to evaluate the need to move to a higher tier. The decision to move an assessment to a higher tier of complexity should result in a revised workplan and consultation with the Agency. The previous PRA, and its sensitivity analysis, should be included in the revised workplan, along with a point estimate risk assessment based on any data collected as part of a lower tier.

Throughout the process of developing the PRA, the EPA risk assessor and the personnel involved in developing the assessment should have a continuing dialogue to discuss the many decisions and their potential impact on the assessment. This dialogue, along with interim deliverables, will help to ensure that the risk assessment report will meet the needs of the Agency and that any problems are identified and corrected early in the process.

6.2 CHECKLIST FOR REVIEWERS

The exposure pathways and chemicals considered in a PRA should be clearly stated and related to the assessment endpoint. Often, the simplest way of doing this is to use the site conceptual model.

Table 6-1 provides a list of major points that may be used to evaluate the quality of a probabilistic assessment. This is not an exhaustive list. The ultimate judgment of the acceptability of a PRA is the responsibility of the regional EPA personnel.

The issues that a reviewer should focus on may be different for each assessment. The workplan and the assessment should address each of the items on the checklist, but the workplan may include additional items. The reviewer is responsible for ensuring that the workplan and the assessment are complete and of sufficient quality to support a risk management decision under the NCP.

The report should include a discussion of the results of assessment and how they relate to the point estimate of risk and hazard. A clear and concise description of what the results mean is a necessary part of each report. This description is best provided in an executive summary near the beginning of the document.

6.3 INTERNAL AND EXTERNAL REVIEW

There are two levels of review that may be appropriate for a PRA. If an EPA reviewer feels they need help with a review, other EPA personnel may be contacted formally or informally to provide additional review capabilities. EPA personnel should also review the draft workplan for PRA to evaluate the appropriateness and consistency with Agency guidance. If EPA personnel are contacted early in the risk assessment process, the review can occur in a more productive and timely manner.

When the issues at a particular site are complex or contentious, EPA reviewers may also wish to obtain the services of outside experts for peer review. According to EPA's Peer Review Policy Statement dated June 7, 1994 (U.S. EPA, 1994), "Major scientifically and technically based work products related to Agency decisions normally should be peer-reviewed." External peer review should be considered when allocating resources for a PRA. EPA reviewers generally should select external peer reviewers who possess no bias or agenda regarding the process or methods of PRA.

Table 6-1. Example of a Generic Checklist [2 pages]

Focal Point	U	Evaluation Criterion
Objectives and Purpose		
Assessment Endpoints	**U**	Are the human health and/or ecological assessment endpoints clearly stated and consistent with the workplan?
Benefits	**U**	Are the rationale for, and benefits of, performing the PRA clearly stated and consistent with the workplan?
Site Conceptual Model	**U**	Is there a description or graphic representation of the receptors and pathways considered in the assessment? Has the PRA addressed each of the pathways for completeness (e.g., sources, release mechanisms, transport media, route of entry, receptor)?
Separation of Variability and Uncertainty	**U**	What is the modeling strategy for separating variability and uncertainty in the PRA? Is this strategy consistent with the assessment endpoints?
Input Distributions and Assumptions		
Variability and Uncertainty	**U**	Is there a clear distinction and segregation of distributions intended to represent variability from distributions intended to represent uncertainty?
Data sources	**U**	Are the data or analysis sources used in developing or selecting the input distributions documented and appropriate for the site?
Distribution Forms	**U**	Are the analyses used in selecting the form of the distribution adequately documented (e.g., understandable and repeatable by a third party?)
Distribution Parameters	**U**	Are the analyses used to estimate the distribution parameters adequately documented?
Distribution Tails	**U**	Do the estimation methods precisely depict the tails of the input distributions; how was this evaluated? Is there sufficient information to depict tails for empirical distributions? Are these estimated as exponential tails with bounding values?
Truncations	**U**	Are any input distributions truncated? Do these truncations make sense? Should truncations be applied to any of the distributions?
Concentration Term	**U**	Is the derivation of a point estimate or distribution for the concentration term adequately documented? Is sufficient information provided to enable the reviewer to recreate the concentration term?
Variable Correlations	**U**	Have variable independence and correlations been addressed? Has the methodology for representing variable correlations in the model been documented and is it reasonable in terms of the variables, the site, and the statistical approach?
Time Step	**U**	Has the basis for the time step used in the model been documented? Is a single time step used, or do variables have different time steps? Are the time steps conceptually reasonable for the variables, for the site? Has the time step been evaluated in the sensitivity analysis?
Sensitivity Analysis	**U**	Has a sensitivity analysis been conducted? Are the methods used in the analysis statistically valid? What did the analysis reveal about uncertainties in the assessment and the relative contributions of input variables to uncertainty?

Focal Point	U	Evaluation Criterion
Model Structure and Computational Mechanics		
Flow Chart	**U**	Is a diagram of the computational sequence provided so that the pathways of inputs and outputs and data capture can be understood and easily communicated?
1-D/2-D MCA	**U**	Is a 1-D MCA or 2-D MCA being implemented in the PRA? What is represented by either or both dimensions?
Algorithms	**U**	Are all algorithms used in the model documented in adequate detail to recreate the analysis?
Integration	**U**	Are the algorithms used in numerical integration identified and documented?
Dimensional Analysis	**U**	Has a unit analysis been conducted to ensure that all equations balance dimensionally?
Random Number Generation	**U**	What random number generator is used in model computations? Is it robust enough? What reseeding approach is used to minimize repeated sequences?
Results of Modeling		
Completeness	**U**	Are all the exposure routes identified in the conceptual site model and workplan addressed in the model results? Has the PRA fulfilled the objectives and satisfied the purpose stated in the workplan?
Point Estimate Calculation	**U**	Has a point estimate calculation, using mean or median values of the input distributions, been performed? How do these results compare with the central tendencies calculated with the probabilistic model? How do the RME estimates compare? Have the similarities or differences between risk estimates from the point estimate and probabilistic approaches been adequately addressed?
Stability of Output Tails	**U**	Has the stability of the high end tail of the risk distribution been adequately evaluated? How stable are the estimated tails (in quantitative terms?) Is this level of stability adequate to support the risk management decisions that the model is intended to support?
Significant Figures	**U**	Is the number of significant figures used in the output reasonable and consistent with model uncertainty?
Limitations	**U**	Are the strengths and weaknesses of the PRA methodology and limitations of the results for decision making clearly presented?
Clarity	**U**	Are the results and conclusions clearly presented and consistent with model output (e.g., CTE and RME identified in the Executive Summary along with discussion of uncertainty)?
Graphics	**U**	Are the graphics included that show both the risk distribution and PRA results (e.g., CTE and RME risk)?

6.4 FOCAL POINTS FOR PRA REVIEW

In reviewing a PRA, it is recommended that a systematic approach be adopted to ensure that all key technical elements of the PRA are evaluated and potential weaknesses are identified. A review check list can facilitate this process and promote consistency in the reviews of PRAs. Such a list can be developed from EPA's guiding principles (U.S. EPA, 1997) and other reviews on the subject of PRA quality review (e.g., Burmaster and Anderson, 1994).

In general, the review of a PRA can be organized into four focal points listed in Exhibit 6-2. PRAs can vary in complexity, from relatively simple to very complicated; thus, the review strategy may need to be customized for specific sites.

EXHIBIT 6-2

FOCAL POINTS FOR PRA REVIEW

1. Clarity of and conformation to objectives;
2. Scientific basis and documentation of input distributions and assumptions;
3. Model structure and computational mechanics; and
4. Results, including, limitations, reasonableness, and clarity of documentation.

Experience with PRA methodology can be invaluable in the review process. Risk assessors should become familiar with PRA methodology. Additional experience with PRA methodology can be gained by obtaining actual site data and performing a probabilistic assessment using this data, and by reviewing and discussing other probabilistic risk assessments. Examples of PRAs can be found in the scientific literature, as well as in the administrative records of Superfund sites, which are available at EPA regional offices.

6.5 ADDITIONAL INFORMATION

Uncertainty: A Guide to dealing with Uncertainty in Quantitative Risk and Policy Analysis (Morgan and Henrion, 1990) and Probabilistic Techniques in Exposure Assessment (Cullen and Frey, 1999) provide excellent philosophical and practical treatises on probabilistic risk assessment. These works are highly recommended to risk assessors who wish to know more about probabilistic risk assessment. The *Summary Report for the Workshop on Monte Carlo Analysis* (U.S. EPA, 1996) and the *Summary Report for the Workshop on Selecting Input Distributions for Probabilistic Assessments* (U.S. EPA, 1999) are other sources of information to learn more about PRA.

REFERENCES FOR CHAPTER 6

Burmaster, D.E. and P.D. Anderson. 1994. Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessment. *Risk Anal*. 14(4): 477–481.

Cullen, A.C. and Frey, H.C. 1999. Probabilistic Techniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. Plenum Press.

Morgan, G.M. and M. Henrion, 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge University Press, New York.

U.S. EPA. 1994. Memorandum from Deputy Administrator Carol Browner on *Peer Review and Peer Involvement at the U.S. Environmental Protection Agency*. June 7, 1994.

U.S. EPA. 1996. *Summary Report for the Workshop on Monte Carlo Analysis.* Risk Assessment Forum. EPA/630/R–96/010.

U.S. EPA. 1997. Memorandum from Deputy Administrator Fred Hansen on the *Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment*, and *Guiding Principles for Monte Carlo Analysis.* Office of Research and Development, Washington, DC. EPA/630/R–97/001. May 15, 1997.

U.S. EPA. 1999. *Summary Report for the Workshop on Selecting Input Distributions for Probabilistic Risk Assessment*. Risk Assessment Forum. EPA/630/R-98/004.

CHAPTER 7

USING PRA TO CALCULATE PRELIMINARY REMEDIATION GOALS

7.0 INTRODUCTION

Chemical-specific preliminary remediation goals (PRGs) are concentration goals for individual chemicals that exist in contaminated media under a specific land use (see Exhibit 7-1). PRGs may be either: (1) concentrations based on applicable or relevant and appropriate requirements (ARARs); or (2) concentrations determined from exposure scenarios evaluated prior to or based on results of the risk assessment.

The National Oil and Hazardous Substances Pollution Contingency Plan (NCP) (U.S. EPA, 1990) outlines two major objectives of risk assessment: (1) to determine if remediation is necessary (i.e., *Is there unacceptable risk at the site?*); and (2) if so, to determine chemical concentrations associated with levels of risk that will be adequately protective of human health and the environment. Chapter 7 deals with this second objective within the framework of PRA.

PRGs are specific to both the medium considered and the land use (U.S. EPA, 1991). PRGs are used in the development of remedial alternatives. Risk-based PRGs are used when ARARs do not exist or in situations when ARARs are not sufficiently health-protective. To develop a PRG based on a point estimate of risk, the standard risk equation is algebraically rearranged to solve for a chemical-specific concentration goal that corresponds to a 10^{-6} risk level for carcinogens or an HI of 1 for noncarcinogens. For PRA, development of PRGs is somewhat more involved.

EXHIBIT 7-1

TERMINOLOGY FOR CHEMICAL CONCENTRATIONS

Preliminary Remediation Goal (PRG) - health-based chemical concentration in an environmental media associated with a particular exposure scenario. PRGs may be developed based on generic exposure scenarios prior to the baseline risk assessment (U.S. EPA, 1991).

Site-specific PRG - health-based chemical concentration based on exposure scenarios in the baseline risk assessment. Generally calculated for a variety of exposure scenarios.

Remediation Goals (RG) - health-based chemical concentration in an environmental medium chosen by the risk manager as appropriate for a likely land use scenario.

Cleanup Level - chemical concentration chosen by the risk manager after considering both RGs and the nine selection-of-remedy criteria of the NCP (U.S. EPA, 1990; 40CFR 300.430(e)(9)(iii)). Also referred to as Final Remediation Levels (U.S. EPA, 1991), chemical-specific cleanup levels are documented in the Record of Decision (ROD). A cleanup level may differ from a PRG because risk managers may consider various uncertainties in the risk estimate, the technical feasibility of achieving the PRG, and the nine criteria outlined in the NCP.

This chapter first reviews the issues associated with deriving PRGs from both point estimate risk assessment and PRA. Next, two distinct methods for developing PRGs from PRA are discussed. Finally, the details of developing PRGs for various environmental media are presented. A clear understanding of the risk assessment goals is needed when using probabilistic methods to develop PRGs. Examples are presented in Appendix D.

1

EXHIBIT 7-2

DEFINITIONS FOR CHAPTER 7

95% UCL for mean - The 95 percent upper confidence limit for a mean of a population is defined as a value that, when repeatedly calculated for randomly drawn subsets of size (*n*), equals or exceeds the true population mean 95 percent of the time. Although the 95% UCL provides a conservative estimate of the mean, it should not be confused with a 95th percentile. As sample size increases, the difference between the UCL for the mean and the true mean decreases, while the 95th percentile of the distribution remains relatively unchanged.

95th Percentile - The number in a distribution such that 95 percent of the values are less than the number and 5 percent are greater.

ARARs - Applicable or relevant and appropriate requirements. The NCP states that ARARs shall be considered in determining remediation goals. If an ARAR meets the requirements of the NCP (U.S. EPA, 1990) for protectiveness, it may be selected as a site-specific cleanup level.

Backcalculation - A method of calculating a PRG that involves algebraic rearrangement of the risk equation to solve for concentration as a function of risk.

Exposure Point Concentration (EPC) - The contaminant concentration within an exposure unit to which receptors are exposed. Estimates of the EPC represent the concentration term used in exposure assessment.

Exposure Unit - For Superfund risk assessment, the geographical area about which a receptor moves and contacts a contaminated medium during the period of the exposure duration.

Iterative Truncation - A method of calculating a PRG that involves developing an expression for the concentration term in which high-end values are "truncated" to reduce the maximum concentration, and calculating risks associated with the reduced concentration. The method may be repeated with consecutively lower truncation limits until risk is acceptable. Iterative truncation methods avoid difficulties associated with applying Monte Carlo analysis to a backcalculation (see above).

Maximum Detected Concentration (MDC) - The maximum concentration detected in a sample (i.e., a set of measurements).

Preliminary Remediation Goal (PRG) - health-based chemical concentration in an environmental media associated with a particular exposure scenario. PRGs may be developed based on applicable or relevant and appropriate requirements (ARARs), or exposure scenarios evaluated prior to or as a result of the baseline risk assessment.

Remediation Action Level (RAL) - A concentration such that remediation of all concentrations above this level in an exposure unit will result in the 95% UCL being reduced to a level that does not pose an unacceptable risk to an individual experiencing random exposures. The RAL will depend on the mean, variance, and sample size of the concentrations within an exposure unit as well as considerations of short term effects of the chemicals of concern.

True Mean Concentration - The actual average concentration in an exposure unit. Even with extensive sampling, the true mean cannot be known. Only an estimate of the true mean is possible. With more samples, the estimate of the mean should be closer to the true mean.

2

7.1 WHEN TO USE PRA FOR DEVELOPING PRGS

In general, PRA methodology is appropriate for developing PRGs in cases where PRA is also used to estimate site-specific risks. Embedded in a site-specific PRG are all of the exposure assumptions and toxicity metrics used in the risk assessment.

- *EPA will generally not accept a PRG with probabilistic methods if probabilistic methods were not used in the risk assessment. The tiered approach for PRA should be followed.*

As indicated by the tiered approach recommended by this guidance for conducting a PRA (Chapter 1, Figure 1-4), a point estimate risk assessment (Tier 1) should always accompany a PRA because it provides information that may be used for risk-management decision making. Using the tiered approach, a risk assessor can determine the appropriate level of complexity that is supported by the available information on exposure and toxicity. These recommendations apply to both the risk assessment and the calculations used to select a PRG. If a PRA is performed, the risk manager may select a PRG based on either the PRA or the point estimate risk assessment. The ultimate cleanup level will be based on the PRG selected by the risk manager and the nine remedy selection criteria (U.S. EPA, 1990; 40CFR 300.430(e)(9)(iii)).

7.2 GENERAL CONCEPTS ABOUT PRGS

PRGs developed from point estimate and probabilistic risk assessments will be discussed in this section to compare and contrast the two approaches.

For both types of risk assessments, the PRG is analogous to the method used to characterize the exposure point concentration (EPC) in the risk equation. For example, if the 95% upper confidence limit for the mean (95% UCL) is used as an EPC, then the PRG should represent a 95% UCL (Bowers et al., 1996). Similarly, if a maximum detected concentration (MDC) is used as an EPC, then the PRG represents the maximum concentration associated with a health-protective level of risk. If uncertainty in the EPC is characterized by a probability distribution rather than a point estimate, then the PRG will correspond to an estimated level of confidence in a target risk estimate. The calculation of PRGs from point estimates of RME risk has been discussed in a number of EPA guidance documents (RAGS I, Part B; Region 3 Risk-based Concentration (RBC) tables; Region 9 PRG tables; U.S. EPA, 1995). The calculation of PRGs from distributions of variability and uncertainty in risk is discussed in this chapter.

Calculation of the EPC generally requires knowledge of not only chemical concentration measurements within the exposure unit but also the receptor's behavior, which determines the exposure unit. For all exposure scenarios, an exposure unit should represent a geographical area in which the majority of the activities of individual receptors occur. Both PRGs and EPCs should be considered with the concept of the exposure unit in mind, and risk assessments should be directed at the area that represents the exposure unit for the target receptor. If an individual is randomly exposed within the same exposure unit over a long period of time, the most appropriate metric for the EPC is the true arithmetic mean concentration with the exposure unit. However, if the exposure is not truly random, a weighted mean may be preferable over the arithmetic mean as a metric for either the EPC or the subsequent PRG. Approaches for weighting the sample data according to site-specific exposure patterns generally should be developed in consultation with EPA regional risk assessors. In addition, for some sites, it may be

appropriate to distinguish between the exposure unit and the remediation unit, which is generally a subset of the exposure unit where remediation may take place. For simplicity in this guidance, it is assumed that the exposure unit and remediation unit represent the same area. Risk assessors may refer to EPA's Soil Screening Guidance (U.S. EPA, 1996) for further guidance on the exposure unit.

In point estimate risk assessments, the 95% UCL is generally used as the EPC to account for the uncertainty associated with estimating the true mean concentration within an exposure unit. When the sample size is small and the variance is large, the 95% UCL may exceed the maximum detected concentration (MDC). In such a case, the MDC is generally used to estimate the EPC, although the true mean may still be higher than this maximum value (U.S. EPA, 1992). For poorly characterized sites, there may be considerable uncertainty that site remediation will be sufficient to reduce the 95% UCL to a health-protective level.

The following paragraphs refer to pre-remediation and post-remediation 95% UCLs. The pre-remediation 95% UCL should generally be determined based on existing site sampling at the time of the Remedial Investigation (RI) prior to remediation. The post-remediation 95% UCL is a prediction of site conditions after remediation. The post-remediation value can be determined by substituting the nondetect level (generally, half the laboratory reporting limit) for some of the high measurements in the concentration sample and recalculating the 95% UCL. The post-remediation value represents expected future conditions at the site after remediation has occurred.

If the risk associated with a pre-remediation 95% UCL exceeds a protective level of concern, a risk assessor may determine how much lower the concentrations would need to be in order to achieve an acceptable risk level. In other words, a maximum allowable concentration may be determined, which this guidance will refer to as a *remediation action level* (RAL) for purposes of site remediation. The challenge for the risk assessor is to identify the RAL that yields an acceptable post-remediation EPC. As discussed above, if the pre-remediation EPC is represented by the 95% UCL, then generally the post-remediation EPC should also be represented by the 95% UCL, rather than the maximum post-remediation concentration. Lowering the RAL effectively lowers the estimated post-remediation 95% UCL. The appropriate RAL depends on the concentrations measured within an exposure unit, and the method used to estimate the 95% UCL. For example, if the 95% UCL is estimated from concentrations fit to a lognormal distribution (U.S. EPA, 1992), the appropriate RAL will depend on the sample size, mean, and variance of the concentrations. If non-parametric methods are used to estimate the 95% UCL (U.S. EPA, 1997), the RAL will depend on the sample size and the skewness of the concentrations. Furthermore, for soil contamination scenarios, assumptions regarding the extent of the exposure unit requiring remediation, as well as the concentrations in backfill, may play a role in determining an appropriate RAL. For cases of poor site characterization, the level will be lower than that of well-characterized sites. In general, as uncertainty increases, a greater remediation effort may be needed to achieve the same health-protective level.

In PRA, either a point estimate or probability distribution may be used to characterize uncertainty in the arithmetic mean concentration (see Chapter 4). The tiered approach described in Chapter 1 (see Figure 1-4) can be followed to determine if a probability distribution should be developed. A distribution for uncertainty in the post-remediation EPC will depend on the arithmetic mean, variance, and number of measurements in the sample, as well as the remediation action level. The appropriate RAL is generally a level that reduces the EPC sufficiently to yield an acceptable risk at the percentile of the risk distribution chosen to represent the RME individual. Multiple simulations of the PRA model will generally be needed to identify a health-protective RAL. (See Section 7.3.2 on Iterative Methods).

To ensure that actual cleanup based on a RAL is protective, post-remedial confirmation sampling is generally required. This step in the risk management process is emphasized further in Section 7.4 on Measurement of Attainment.

EPA has advocated a Data Quality Objectives (DQO) process to determine the type, quantity, and quality of data needed to make defensible decisions (U.S. EPA, 1993). This guidance also advocates the DQO process to provide measures of confidence that EPCs and PRGs are representative and appropriate for characterizing site conditions.

In PRA, the risk manager will be interested in developing a PRG that accounts for both variability and uncertainty. Ultimately, the risk management goal is to protect a large percentage of the population with reasonable confidence, thus remaining in accord with the NCP (U.S. EPA, 1990). To the extent practicable, variability and uncertainty are treated separately in PRA. By characterizing variability, risk managers can identify a risk estimate from a probability distribution for risk (i.e., the RME range discussed in Section 4.4). By characterizing uncertainty, the confidence in the risk estimate can be evaluated.

It is appropriate to distinguish between variability and uncertainty in the risk estimate from PRA. While variability in the EPC may be characterized for scenarios evaluated using microexposure event (MEE) analysis (see Appendix E, Section E.2), typically uncertainty in the EPC is of greatest interest. In a 1-D MCA, probability distributions will be used to characterize variability *or* uncertainty, but not both. It is anticipated that for human health risk assessments, probability distributions will typically characterize variability in any exposure variable except the EPC, which will be represented by a point estimate for uncertainty in the mean concentration (e.g., the 95% UCL or MDC). In contrast, for ecological risk assessments (see Chapter 5), probability distributions for uncertainty in toxicity and exposure variables (including the EPC), are typically of greatest interest in a 1-D MCA.

In a 2-D MCA, probability distributions may be developed to characterize both variability and uncertainty (see Section 4.3 and Appendix E). Such analyses may be used to determine the relative contribution of uncertainty in the EPC to uncertainty in the risk estimate (see Chapter 2 on Sensitivity Analysis).

7.3 METHODS FOR CALCULATING PRGS

There are two primary methods for developing PRGs from PRA: backcalculation and iterative methods using predicted post-remediation concentrations. Iterative truncation of the pre-remediation concentrations is a common iterative method..

L *EPA recommends that PRGs be developed using iterative methods.*

In some unique cases, given constraints on the shapes of the input distributions, an analytical solution for the output distribution may be used. In this case, a Monte Carlo simulation would not be needed.

7.3.1 BACKCALCULATION

Traditionally, risk is calculated as a function of multiple exposure variables, including the concentration term, and toxicity value (Equation 7-1). If one or more of the exposure variables is described by a PDF, a Monte Carlo simulation will yield a distribution for risk (see Chapter 1). Backcalculation methods can be envisioned as holding the risk level at a constant and acceptable level (e.g., risk equal to 10^{-6} or HI equal to 1) and then algebraically reversing the risk equation to solve for the concentration term (Equation 7-2). A Monte Carlo simulation using Equation 7-2 will yield a distribution of concentrations that reflects the combination of distributions from all other exposure variables.

$$\frac{C \cdot CR \cdot EF \cdot ED}{BW \cdot AT} \cdot Toxicity = Risk$$

$$C \cdot V = Risk \qquad \text{Equation 7-1}$$

$$C = Risk \cdot V^{-1} \qquad \text{Equation 7-2}$$

where the toxicity metric is based on the cancer slope factor (CSF) or the reference dose (1/RfD) for the chemical in the exposure medium, *C* is the concentration term, and *V* represents the algebraic combination of the toxicity metric with all exposure variables except *C*.

Presently, there is considerable debate in the field of probabilistic risk assessment about the appropriateness of backcalculation. Some investigators have claimed that using a point estimate for risk and deriving a distribution for concentration via backcalculation is an appropriate technique (Stern, 1994; Burmaster et al., 1995; Bowers, 1997). In this view, the percentiles of the concentration distribution resulting from backcalculation mirror the percentiles of the risk distribution. Said another way, if the PRG is identified from the (1.0 - α) percentile of the cumulative distribution for risk (forward calculation), where α is p(Risk > target risk level), then PRG corresponds with the (α) percentile of the cumulative distribution for concentration (backcalculation).

For example, assume *C* represents a 95% UCL of 100 ppm, α = 0.03, and the target *Risk* is $1x10^{-6}$. Equation 7-1 yields a distribution for risk in which there is a 97 percent (i.e., 1 - α) probability that risk is no greater than $1x10^{-6}$ given a 95% UCL of 100 ppm. Similarly, Equation 7-2 yields a distribution for concentration in which there is a 3 percent probability that a value less than the 100 ppm is needed to yield a target risk of $1x10^{-6}$. This correspondence between Equations 7-1 and 7-2 is reliable only when both the concentration term and the risk are characterized by point estimates.

The correspondence between the output distributions from the forward and backward equations is maintained even for more complex models involving multiple exposure media. Equation 7-3 is an extension of Equation 7-2 whereby PRG is calculated for one exposure medium (C_1), but risk is a function of exposure to a contaminant in three media. Note that the concentration terms for the additional exposure media (C_2, C_3) must be expressed as point estimates rather than distributions (for reasons described below).

$$C_1 = \left[Risk \cdot (C_2 \cdot V_2 + C_3 \cdot V_3) \right] \cdot V^{-1} \qquad \text{Equation 7-3}$$

There are significant limitations in the use of backcalculation in PRA (Ferson, 1996). When either concentration or risk is characterized by a distribution, instead of point estimates, terms in the risk equation can not be algebraically rearranged (Burmaster et al., 1995; Ferson, 1996). The difficulty for PRA arises because each risk estimate from a Monte Carlo analysis that uses the familiar "forward-facing" risk equation represents a combination of random values selected from the input distributions. Therefore, the output can be considered conditional on all of the inputs. Rearranging the risk equation does not maintain the same conditional probabilities; therefore, the distribution for risk estimated as a function of the distribution for concentration in Equation 7-1 does not return the same distribution for concentration when applied in Equation 7-2. In addition, the correlations of the inputs to the risk equation may not be known, and the inputs may not be independent.

Another difficulty with the backcalculation approach is that it provides a limited characterization of the post-remediation concentrations. Concentrations remaining on site after remedial activities may need to satisfy more than one regulatory constraint. For example, the average or 95% UCL concentration within an exposure unit may need to be less than a level based on chronic toxicity or cancer and simultaneously, the RAL concentration may need be less than a level based on acute toxicity.

Because the backcalculation approach does not specify a RAL, additional calculations would be needed to evaluate multiple health endpoints. Backcalculation methods by themselves do not provide sufficient information to address concerns about acute toxicity. In general, acute toxic effects from high exposure (e.g., children with pica behavior for soil) at the RAL should be considered for all chemicals. Acute toxicity is discussed further in Section 7.3.2.

In general, while backcalculation methods may be recommended for some sites, risk assessors should be familiar with their limitations. Because of these limitations, this guidance recommends iterative forward calculations as a method for calculating PRGs. Iterative approaches are generally more flexible than backcalculations and can provide more information for making remedial decisions.

7.3.2 ITERATIVE METHODS

EPA recommends iterative simulations as a general approach for calculating PRGs from probabilistic risk assessments. The simplest method involves calculating risk with the "forward-facing" equation (Equation 7-1) a number of times (iteratively). With each iteration, a different estimate of the 95% UCL is used until the risks are acceptable. This iterative method has also been called the "repeated runs" method. Computer software is commercially available to facilitate iterative calculations (e.g., EXCEL's Solver Routine), although often iterative calculations can be performed by simply using a systematic trial-and-error approach. This method will yield the same result as a backcalculation approach, and provides a probability distribution for risk that can be used in the risk communication process (Chapter 8). In addition, unlike the backcalculation approach, a probability distribution can be used to characterize EPC (see Appendix E).

7.3.2.1 ITERATIVE TRUNCATION

The iterative truncation method involves truncating the higher values in the sample of concentration measurements, developing a new "truncated", or post-remediation concentration term, and calculating risk with the "forward-facing" equation (Equation 7-1) to determine if the risks are acceptable. With each iteration, a different RAL concentration is specified and the EPC is recalculated until the risk distribution yields risk estimates at or below a level of concern.

Iterative truncation may not be appropriate for every site. Criteria for this method are presented in Exhibit 7-3. To some extent, potential limitations highlighted by these criteria may also introduce uncertainty in point estimates of the EPC that are based on the 95% UCL. For example, if the sample size is too small and the data are highly skewed, the 95% UCL may exceed the maximum detected concentration. Similarly, certain environmental sampling approaches may introduce bias in the estimate of the 95% UCL. Of course, both point estimate and probabilistic methods are sensitive to poor site characterization.

The iterative truncation method is easiest to think about with regard to soil cleanup when contaminated soil is removed and replaced with clean fill dirt. This replacement would reduce both the mean and 95% UCL. In the example given in Chapter 4 for the hypothetical wood treatment facility (see Section 4.4.8), an acceptable 95% UCL was defined as an EPC that yielded a 97th percentile risk of no greater than $1x10^{-5}$. When a sufficient portion of the areas with high concentrations has been removed, the resulting post-remediation risks will fall within the acceptable range. Under these assumptions, the truncation point (i.e., the concentration above which removal occurred) is a true remediation action level. Post-remediation sampling can confirm that the concentration term used in the risk assessment adequately represents conditions following site remediation (see Section 7.4 on Measurement of Attainment).

Because PRGs determined from iterative truncation are easiest to visualize for contaminated soil, the discussion in the next several paragraphs will be based on soil

EXHIBIT 7-3

CRITERIA FOR ITERATIVE TRUNCATION

1. **Sample size (*n*) is sufficient.** Small sample sizes may preclude developing reliable estimates of uncertainty in the concentration term. The risk assessor should determine that *n* is sufficient to produce a distribution for concentration that is representative of site conditions.

2. **Concentration distribution is not highly skewed.** A highly skewed distribution may yield unreliable estimates of uncertainty, especially for small sample sizes.

3. **Sampling design yields a representative distribution of measurements within the exposure unit.** Simple random sampling may fail to represent a patchy spatial distribution of contaminants. Similarly, hotspot (e.g., cluster) sampling may fail to represent random movement of receptors. To evaluate potential biases in sampling, analyses with both standard statistical methods and geostatistical methods may be required.

4. **Assumptions about the post-remedial distribution of concentration are reasonable.** If these assumptions are shown to be incorrect by subsequent sampling events, the process for developing a PRG may need to be repeated and additional remedial activities may be required.

contamination. Groundwater and other media are considered in later sections. Figure 7-1 shows an example of iterative truncation, assuming pre-remediation and post-remediation concentrations in soil can be adequately represented by lognormal distributions. Table 7-1 summarizes the statistics that are relevant to deriving a PRG at the appropriate risk level. The 95% UCL for the pre-remediation distribution of concentrations (Curve A) is 1650 ppm. A RAL of 1400 ppm is needed to achieve a PRG (i.e., 95% UCL) of 900 ppm, as shown by Curve B. A RAL of 1150 ppm is needed to achieve a protective PRG of 450 ppm, as shown by Curve C.

Each situation is unique in terms of actual RALs and their associated PRGs. Different RALs would be needed if risk were estimated from a different percentile of the uncertainty distribution for the concentration term (e.g., 80% UCL instead of 95% UCL). In addition, the uncertainty distribution is sensitive to the variability in concentrations. For simplicity, post-remediation concentrations in this example are assumed to be lognormally distributed. The site is also assumed to be sufficiently characterized and all criteria in Exhibit 7-3 are met.

A more complex mixed distribution that reflects both the relatively low concentrations in backfill and the higher post-remediation sample concentrations may more faithfully represent the PRG (Figure 7-2). A single lognormal distribution may not adequately characterize heterogeneity in contamination at a hazardous waste site. In practice, a sample of measured concentrations may be represented by a probability distribution and a group of values below the quantitative detection limit (i.e., nondetects). Generally, metals and other inorganic chemicals will be detected in the entire sample. In the iterative truncation method, post-remediation concentrations can be thought of as a mixed distribution that reflects the weighted combination of the truncated pre-remediation concentrations and the very low concentrations in the soil used as clean backfill. In the absence of site-specific data, the surrogate value for nondetects (e.g., half the laboratory reporting limit) may be used as an estimate of the concentration in clean backfill.

A post-remedial distribution for soil will consist of the lower concentrations occurring in the pre-remedial sampling distribution and a number of nondetect values.

Because of the more complex nature of most mixed distributions (Roeder, 1994), non-parametric methods for calculating the 95% UCL of the arithmetic mean (e.g., bootstrap resampling) may be appropriate (U.S. EPA, 1997). Of course, sampling results from surface soil generally should not be averaged with those from subsurface soil unless appropriate for the specific exposure scenario being considered (e.g., excavation worker).

For a number of chemicals, acute toxic effects from high exposure (e.g., children with pica behavior for soil) at the RAL should also be considered. In the absence of Agency guidance on levels that protect against acute toxicity, a toxicologist should be consulted regarding the RALs. EPA anticipates that future guidance will need to be issued regarding acute toxicity, cleanup strategies, and the RAL.

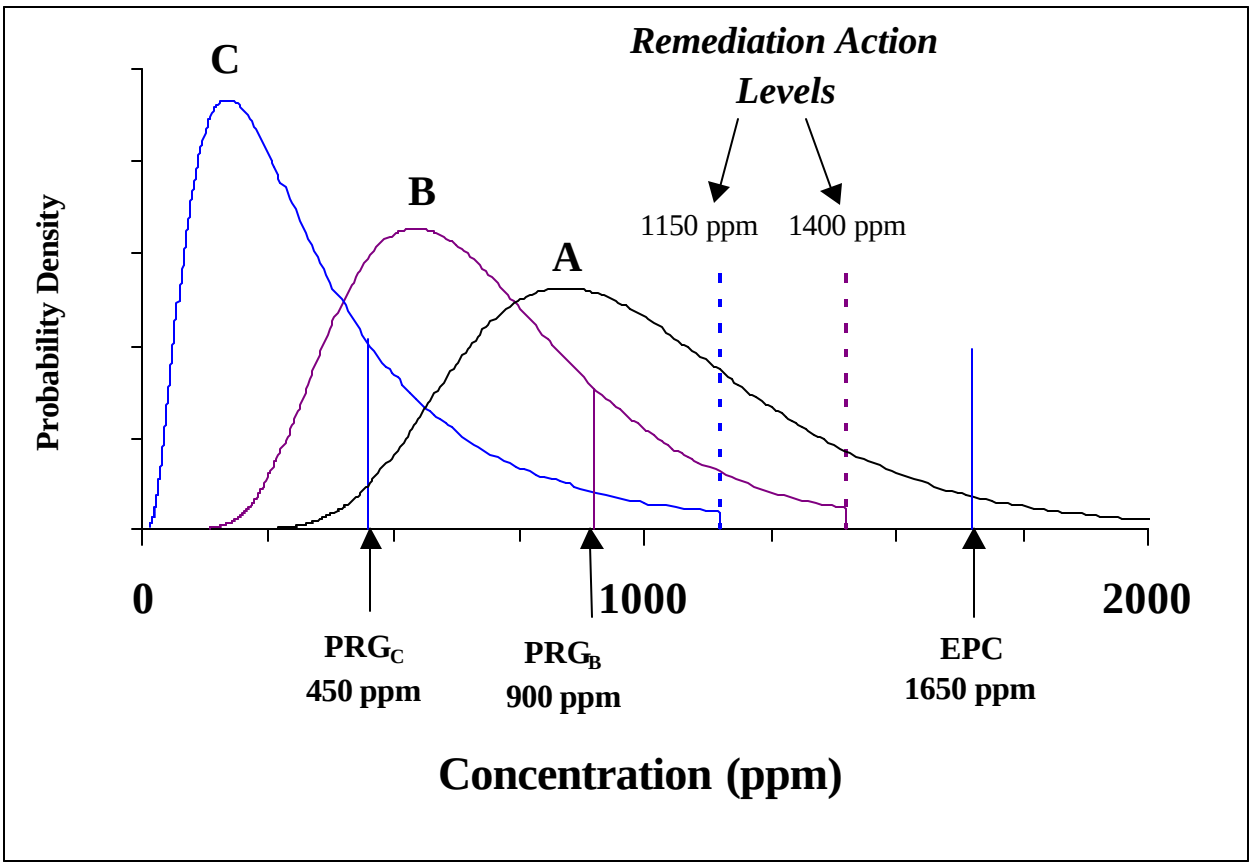


Figure 7-1. Hypothetical examples of distributions of concentration associated with remediation action level (RAL) truncation points. Curve A is the pre-remediation distribution, curve B is the post-remediation distribution with a RAL of 1400 ppm, and curve C is the post-remediation distribution with a RAL of 1150 ppm. The EPC and both PRGs are represented by the 95% UCL for the arithmetic mean of their respective distributions. A RAL of 1400 ppm yields a 95% UCL of 900 ppm (PRG_B), whereas a RAL of 1150 ppm yields a 95% UCL of 450 ppm (PRG_C). Reducing the RAL by 250 ppm results in a 50% reduction in the 95%UCL and corresponding HI (Table 7-1). Each hypothetical distribution represents a truncated lognormal distribution for purposes of illustration; in practice, a more likely scenario would be a mixed, bimodal distribution that represents a combination of the truncated pre-remediation distribution and the relatively low concentration in backfill (Figure 7-2).

Table 7-1. Summary statistics for hypothetical example shown in Figure 7-1.

Lognormal Distribution	Remediation Action Level (ppm)	PRG = 95% UCL (ppm)	Hazard Index (hypothetical)
Curve A	none (+ 4)	1650	3.7
Curve B	1400*	900	2.0
Curve C	1150*	450	1.0

* Imposing a RAL is equivalent to specifying the high-end truncation limit for the lognormal distribution.

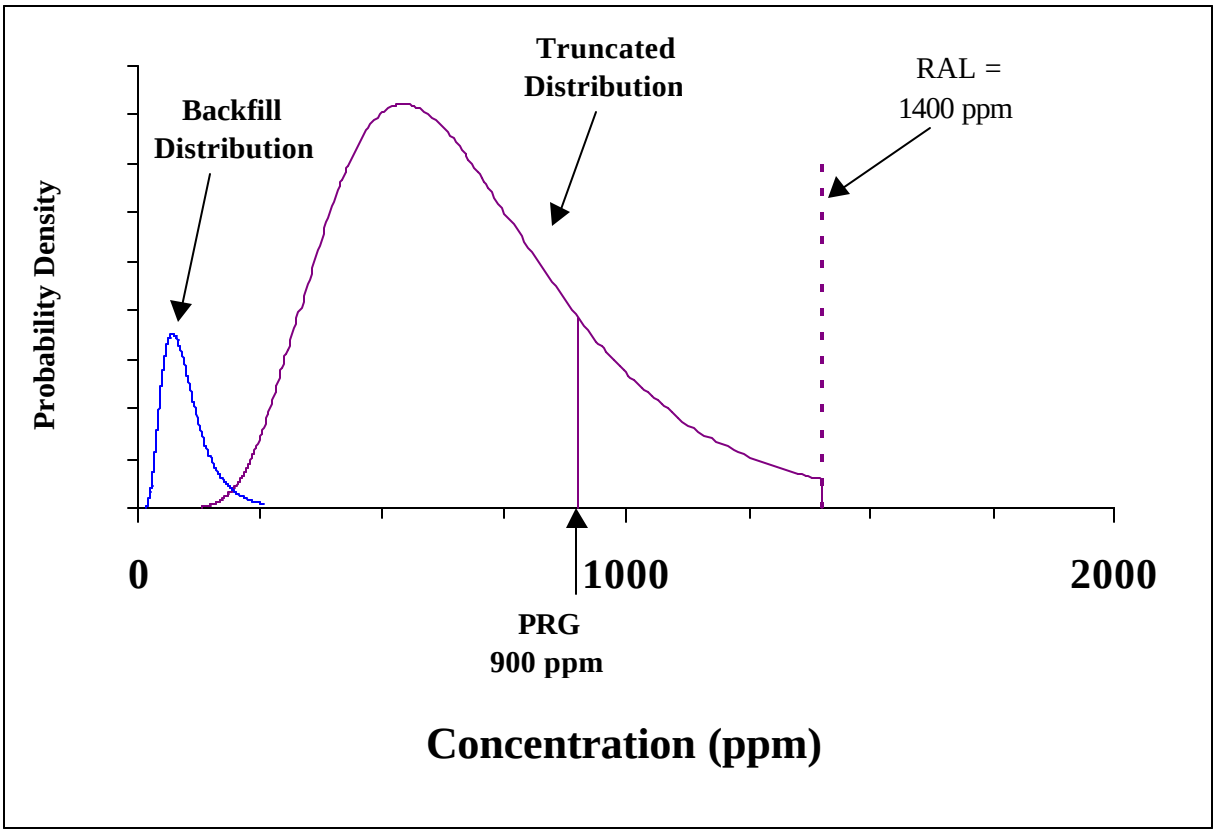


Figure 7-2. Hypothetical example of a mixed, bimodal distribution that represents a combination of the pre-remediation distribution truncated at an remediation action level of 1400 ppm, and a distribution for backfill soil with relatively low concentrations.

7.3.2.2 ITERATIVE TRUNCATION PRGS FOR CONTAMINATED SOIL

At some sites, cleanup levels may be achieved by actual removal or treatment of only a portion of the soil on site. This may reflect a variety of remediation strategies, including the decision to distinguish between a remediation unit and an exposure unit (Section 7.2, General Concepts about PRGs), or a decision to employ multiple remediation technologies. For example, some soil may be left in place while soil in other areas of the site is removed or treated via soil vapor extraction, biological land farming, or some other remedial technology. At some sites, all of the soil may need to be removed or treated. In general, treatment or removal of the more heavily contaminated soil is expected to reduce site risks to below levels of concern. An evaluation of remedial action objectives should be based on confirmation sampling (Section 7.4, Measures of Attainment).

7.3.2.3 ITERATIVE TRUNCATION PRGS FOR GROUNDWATER

For some chemicals encountered at hazardous waste sites, chemical-specific applicable or relevant and appropriate requirements (ARARs) may exist. ARARs are often PRGs and, if they meet the requirements of the NCP (U.S. EPA, 1990) for protectiveness, ARARs may be selected as site-specific cleanup levels. For PRGs and cleanup levels that are based on ARARs, risk-based approaches for developing PRGs may not be appropriate. This is especially true of ARARs that are applied as RALs; site concentrations would be compared to the ARAR, rather than a truncation level needed to achieve an acceptable risk distribution (see Figure 7-1).

- *For groundwater contamination, ARARs should be applied as remediation action levels if they are protective.*

For cases in which an ARAR may not be protective, risk-based levels generally should be developed in accordance with the NCP (U.S. EPA, 1990).

Groundwater is not a static medium. In addition, receptors contact groundwater at specific and unchanging locations (e.g., wellheads). Often, a single well can be considered the exposure unit when considering either the residential or industrial/occupational scenarios. There is considerable uncertainty about the fate and transport of chemicals in groundwater over time. Ideally, the risk assessment would focus on individuals who may be exposed at locations nearest to the center of the contaminant plume, where concentrations are likely to be highest (Freeze and Cherry, 1979; Sposito, Jury, and Gupta, 1986). Point estimates or probability distributions for this concentration term generally should reflect uncertainty in the long-term arithmetic mean concentration rather than variability. A microexposure event (MEE) modeling approach may be appropriate for simulating short-term changes in exposure corresponding to temporal and spatial variability of concentrations at a wellhead (e.g., seasonal fluctuations in the water table). Examples of MEE approaches are discussed in Appendix E. When using an MEE approach, risk assessors may explore different time steps for simulating short term exposure events; however, care should be taken to use an averaging time that is relevant to the toxicity metric (i.e., short term or long term toxicity).

Because of the movement of groundwater and the necessity of sampling the medium at fixed locations, identifying a meaningful RAL needed to achieve a given post-remediation 95% UCL is difficult. For PRA, repeated 1-D MCA simulations of the probabilistic model may be performed using point estimates from a range (or distribution) of values that represent uncertainty in the concentration term. In addition, the distribution representing uncertainty in the mean concentration may be combined with

distributions for uncertainty in other exposure variables using 2-D MCA. A PRG is generally the value that yields an acceptable risk for the percentile of the risk distribution chosen to represent the RME individual.

7.3.2.4 ITERATIVE TRUNCATION PRGS FOR OTHER CONTAMINATED MEDIA

Iterative truncation techniques are generally applied to solid media. At some sites, sediment may be considered a fluid medium. For example, in a harbor with constant boat traffic, sediment may be resuspended by waves, changing tides, or beach erosion.

The development of PRGs for surface water is also difficult with iterative truncation. For fluid media (like groundwater or surface water), repeated simulations of the PRA model can be performed using a range of values to determine a PRG with acceptable risk at the RME percentile.

Cleanup levels for fish tissue are derived in much the same way as for soil and groundwater. The exposure unit, in this case, is the area where the angler population, or ecological receptor population, harvests fish. However, in risk assessments that include a fish ingestion exposure pathway, there is much uncertainty about the true EPC. Concentrations may be affected by many factors, including changes in the fish population and changes in fish preferences, which may be difficult to address in these risk assessments. The choice of fish species consumed by a given individual will also affect the concentration term.

Fish population studies and fate and transport considerations of the contaminants may indicate if and when a fish population will reach a calculated cleanup level. For many sites, it may be difficult to obtain this level of site-specific data within the resource and time constraints of the risk assessment. Although remediation may not immediately reduce contaminant concentrations in fish, the determination of a cleanup level can serve as a target for any future decline in concentrations.

7.4 MEASUREMENT OF ATTAINMENT

The NCP (U.S. EPA, 1990) mandates continued monitoring for groundwater cleanups to ensure attainment of the remedial action objectives. In addition, it is common practice among remedial project managers to conduct confirmation sampling after completing a remedy for soil contamination. For soil contamination, PRGs and RALs are developed as part of the cleanup strategy. However, completion of the remedial action according to this strategy does not necessarily mean that risks within exposure units at the site have been reduced to levels specified in the Record of Decision (ROD). Confirmation sampling following cleanup activities is highly recommended to ensure that any contamination left on site does not pose an unacceptable risk.

- *If confirmation sampling indicates insufficient risk reduction, a more extensive remediation effort will be needed.*

Although deciding to perform post-remedial confirmation sampling is the risk manager's task, this guidance strongly recommends confirmation sampling to determine whether or not the remedial action has been successful. For post-remediation sampling, the Data Quality Objectives process should generally be followed. If the post-remediation risk associated with the confirmation sample indicates risk exceeds a level of concern, then additional remediation may be warranted.

Regarding soil, when only a portion of the contaminated soil within an exposure unit has been removed or treated, both the confirmation sample and the RI samples from unremediated locations should generally be used in a post-remediation risk calculation to ensure that the remedial action objectives have been met. The calculation of post-remediation EPC (using the confirmation sample and a portion of the RI sample) is conceptually similar to the EPC calculation in the baseline risk assessment.

7.5 CHOOSING A CONCENTRATION TO BE A PRG

As with point estimate risk assessments, the goal in PRA generally should be to ensure that the post-remediation risk to the RME individual is at or below an acceptable level (e.g., cancer risk of 10^{-6}, hazard index of 1). For a 1-D MCA, the risk level of concern will typically be selected *a priori* as a percentile of the risk distribution. For example, a risk manager may wish to identify the PRG that corresponds with a 95th percentile risk of $1x10^{-6}$. In general, this RME percentile should be selected from the high-end of the probability distribution for risk, between the 90th and 99.9th percentiles. As discussed in Chapter 1, the 95th percentile of the risk distribution is generally the starting point for selecting a percentile to represent the RME. Chapter 4 (Section 4.4) presents a series of risk assessment and site-specific factors that may guide the risk management decision to select a higher or lower percentile.

PRA may also be used to quantify the uncertainty in the risk estimates corresponding with the percentile range for the RME (i.e., 90th to 99.9th percentile). If there is great uncertainty in the risk estimate corresponding with the percentile chosen to represent the RME, this may support a decision to reject the choice in favor of a percentile higher in the RME range. In contrast, relatively low uncertainty in risk estimates in the upper tail of the risk distribution may support a selection of a lower RME percentile. Quantifying uncertainty in risk estimates provides risk managers with greater flexibility to weigh the various options for remedial action.

Figure 7-3 presents hypothetical results of 2-D MCA simulations in which both variability and uncertainty were characterized for multiple exposure variables. The CDFs for the risk distribution are given along with the 90% confidence interval (CI) for the CDF, as summarized in Table 7-2. For this example, assume $1x10^{-6}$ is the risk level of concern. The top panel shows the 95th percentile risk slightly exceeds $1x10^{-6}$, with a narrow 90% CI [$9.8x10^{-7}$, $1.3x10^{-6}$]. These results suggest that the contaminant concentrations would need to be reduced to achieve an acceptable PRG. The manager may choose a different percentile to represent the RME depending on site-specific information. The risk manager's choice of percentile would be supported by the relatively low uncertainty in the risk estimates, indicated by the narrow confidence interval.

In contrast, the bottom panel shows a cumulative risk distribution with a relatively wide confidence interval. The 90% CI spans an order of magnitude for risks in the tail of the distribution (> 90th percentile), indicating high uncertainty in the risk estimates. For example, the 90% CI for the 95th percentile is [$1.2x10^{-6}$, $1.4x10^{-5}$]. This high uncertainty, together with site-specific information, might support a decision to choose a higher percentile to represent the RME. For example, at the 98th percentile, there is a 90% probability that the risk is between $1.9x10^{-6}$ and $2.1x10^{-5}$. In general, as the percentile of the risk distribution representing the RME is increased, a lower PRG is needed to achieve a risk level of concern with high confidence.

Table 7-2. Results of 2-D MCA simulations shown in Figure 7-3 showing 90% CI [95% LCL, 95% UCL] for selected percentiles of the risk distribution.

Percentile of CDF for Risk	Fig. 7-2 Top Graph		Fig. 7-2 Bottom Graph	
	95% LCL	95% UCL	95% LCL	95% UCL
90th	7.3×10^{-7}	9.5×10^{-7}	9.5×10^{-7}	1.0×10^{-6}
95th	9.8×10^{-7}	1.3×10^{-6}	1.2×10^{-6}	1.4×10^{-5}
96th	1.0×10^{-6}	1.4×10^{-6}	1.4×10^{-6}	1.6×10^{-5}
97th	1.2×10^{-6}	1.6×10^{-6}	1.6×10^{-6}	1.8×10^{-5}
98th	1.4×10^{-6}	1.9×10^{-6}	1.9×10^{-6}	2.1×10^{-5}

7.6 CALCULATING MEDIA-SPECIFIC PRGS FROM PROBABILISTIC ANALYSIS

Whatever medium is considered in the development of PRGs, the risk assessor should keep in mind that the PRG comprises both an average and a RAL that embody aspects of both the spatial distribution of contamination and the movement of the receptor and possibly the contaminated medium within the exposure unit. Table 7-3 provides examples of sources of variability in the concentration term based on both the contamination in selected exposure media (soil, groundwater, and fish) and the receptor.

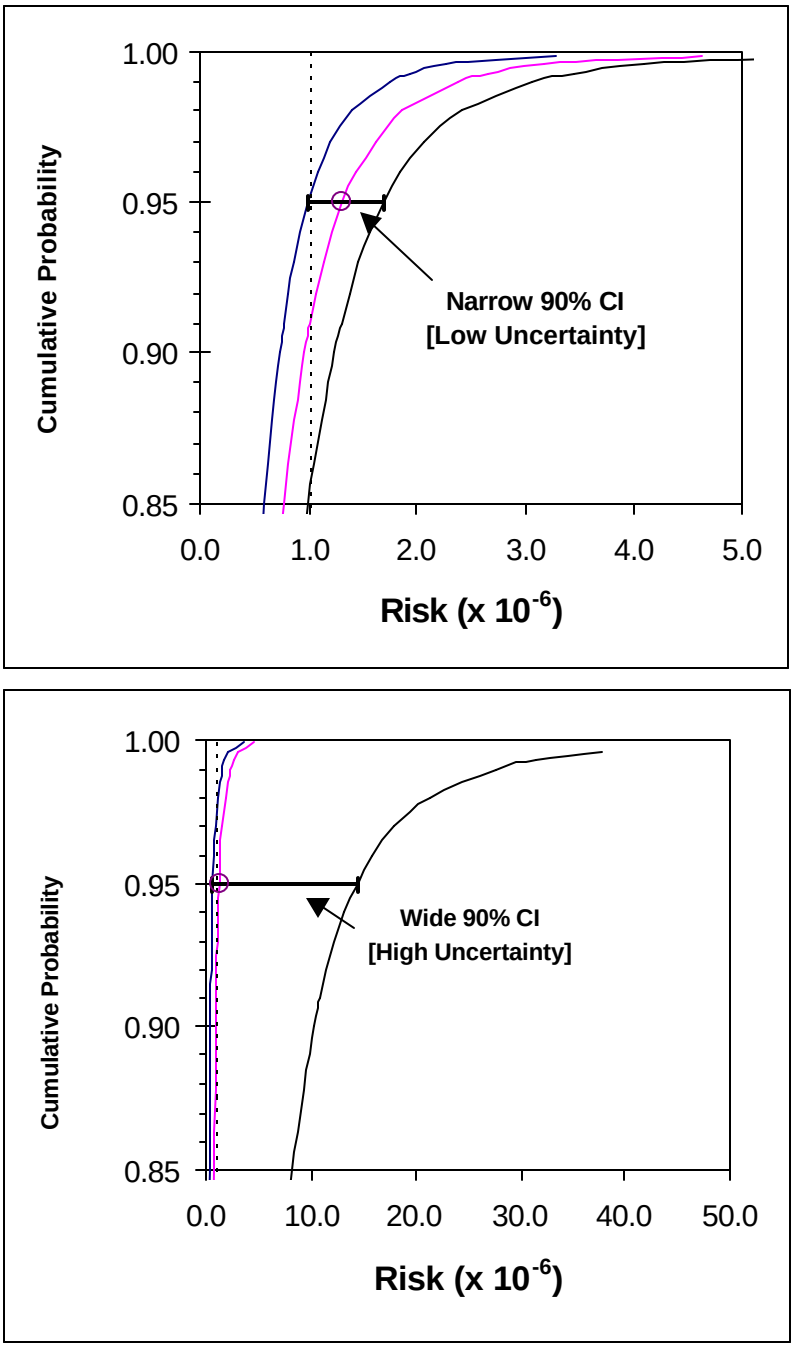


Figure 7-3. Hypothetical examples of 2-D MCA results showing the CDFs for risk. Surrounding each CDF are two confidence limits corresponding with the 5th and 95th percentiles of the distribution for uncertainty. The graphs display only the upper portion of the risk distribution because risk managers should generally use the RME range from the 90th to 99.9th percentiles as a basis for their decisions. This RME range corresponds to the range of 0.9 to 0.999 on the y-axis. In the top panel, the 90% CI for the risk distribution is narrow, which supports the risk manager's choice of the percentile to represent the RME. In the bottom panel, the 90% CI is wide, which may support the decision to choose a higher percentile to represent the RME. In the bottom panel, the lower confidence limit at the 98th percentile is slightly greater than $1x10^{-6}$. Note the scale of risk on the x-axis is much larger in the bottom panel. An RPM may consider similar information for non-cancer hazards.

Table 7-3. Examples of sources of variability in the concentration term for selected exposure media.

Factor		Soil	Groundwater	Fish
Temporal Variability	Contaminant	! none, if contaminant source is inactive ! aerial deposition from ongoing source emissions affected by wind patterns ! degradation over time ! volatilization ! migration to groundwater ! radioactive growth and decay	! seasonal fluctuation in groundwater table ! migration of contaminant plume ! natural attenuation	! seasonal changes in species availability ! bioconcentration ! longterm changes in population dynamics ! fish tissue concentrations linked to temporal variability in water and sediment concentrations ! physical and chemical processes that change concentrations
	Receptor	! non-random movement throughout exposure unit	! none, fixed location at specific wellhead ! changes in well location over time	! dietary preferences for fish species ! cooking practices
Spatial Variability	Contaminant	! heterogeneity over a small area and with depth ! cluster sampling to identify hotspots as opposed to random sampling that more closely estimates the receptor's concentration	! migration of contaminant plume, based on hydrogeology and source emissions (e.g., bulk flow or continuous source)	! migration of fish ! changes in fish population structure
	Receptor	! random movement throughout exposure unit ! daily activity patterns involve multiple exposure units	! none, fixed location at specific wellhead ! changes in well location over time	! change in recreational habits, areas fished

REFERENCES FOR CHAPTER 7

Bowers, T.S., N.S. Shifrin and B.L. Murphy. 1996. Statistical approach to meeting soil cleanup goals. *Environ. Sci. Technol.* 30: 1437-1444.

Bowers, T.S. 1997. Derivation of Remediation Goals Using Monte Carlo Methods. In: Trends in Risk and Remediation. p. 6, Summer, 1997.

Burmaster, D.E., K.J. Lloyd, and K.M. Thompson. 1995. The need for new methods to backcalculate soil cleanup targets in interval and probabilistic cancer risk assessments. *Hum. Ecol. Risk Assess.* 1(1): 89-100.

Ferson, S. 1996. What Monte Carlo methods cannot do. *Hum. Ecol. Risk Assess.* 2: 990-1007.

Freeze, R.A. and J.A. Cherry. 1979. Groundwater. Prentice Hall, Inc., NJ.

Roeder, Kathryn. 1994. A graphical technique for determining the number of components in a mixture of normals. *J. American Stat. Assoc.* 89(426): 487-495.

Sposito, G., W.A. Jury, and V.K. Gupta. 1986. Fundamental problems in the stochastic convection-dispersion model of solute transport in aquifers and field soils. *Water Res. Research.* 22(1): 77-88.

Stern, A.H. 1994. Derivation of a target level of lead in soil at residential sites corresponding to a *de minimis* contribution to blood lead concentration. *Risk Anal.* 14: 1049-1056.

U.S. EPA. 1990. *National Oil and Hazardous Substances Pollution Contingency Plan*. Final Rule. 40 CFR 300: *55 Federal Register*, 8666-8865, Thursday, March 8.

U.S. EPA. 1991. *Risk Assessment Guidance for Superfund* (RAGS), *Volume 1: Human Health Evaluation Manual* (HHEM), *Part B, Development of Risk-Based Preliminary Remediation Goals.*. Office of Emergency and Remedial Response, Washington, DC. EPA/540/R-92/003. NTIS PB92-963333.

U.S. EPA. 1992. *Supplemental Guidance to RAGS: Calculating the Concentration Term*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive 9285.7-081.

U.S. EPA. 1993. *Data Quality Objectives Process for Superfund*. Office of Solid Waste and Emergency Response, Washington, DC. EPA/540/R-93/071.

U.S. EPA. 1995. *Supplemental Guidance to RAGS: Region 4 Bulletins*. #5 Development of Remedial Goal Options.

U.S. EPA. 1996. *Soil Screening Guidance: User's Guide*. Second Edition. Office of Emergency and Remedial Response, Washington, DC. EPA/540/R-96/018.

U.S. EPA. 1997. *The Lognormal Distribution in Environmental Applications*. Office of Research and Development, and Office of Solid Waste and Emergency Response, Washington, DC. EPA/600/R-97/006. December.

CHAPTER 8

COMMUNICATING RISKS AND UNCERTAINTIES IN PROBABILISTIC RISK ASSESSMENTS

8.0 INTRODUCTION

EPA has developed guidance on the role of the community in the Superfund process, *Risk Assessment Guidance for Superfund: Vol. 1 - Human Health Evaluation Manual, Supplement to Part A: Community Involvement in Superfund Risk Assessments* (U.S. EPA, 1999). This supplement to *Risk Assessment Guidance for Superfund* (RAGS) *Vol. 1 Part A* was developed to improve community involvement in the Superfund risk assessment process. It should serve as a primary community involvement resource for risk assessors, along with the *Superfund Community Involvement Handbook and Toolkit* (U.S. EPA, 1998). The *Community Involvement in Superfund Risk Assessments* specifically:

- provides suggestions for how Superfund staff and community members can work together during the early stages of Superfund remedial investigation and later cleanup;
- identifies where, within the framework of the human health risk assessment methodology, community input can augment and improve EPA's estimates of exposure and risk;
- recommends questions the site team (risk assessor, remedial project manager (RPM), and community involvement coordinator) should ask the community; and
- illustrates why community involvement is valuable during the human health risk assessment at Superfund sites.

This chapter provides additional guidance and suggestions to deal with risk communication issues that arise during a probabilistic risk assessment (PRA). It highlights the appropriate level of public involvement and principle risk communication skills needed to effectively communicate risk information from a PRA to various interested parties at a Superfund site. Section 8.2 discusses additional activities for communicating PRA information, while Sections 8.3 and 8.4 provide guidance on specific techniques for communicating information. The success of risk communication efforts will depend on the extent to which the communication strategy addresses the needs of a diverse audience, with different perceptions of risk and uncertainty (Section 8.5), and the degree of trust and credibility that is established from the onset (Section 8.6).

8.1 EARLY INVOLVEMENT AND ENGAGEMENT OF STAKEHOLDERS

There are many stakeholders in a risk assessment (see Chapter 1, Section 1.6). It is generally important to *involve and engage all stakeholders early* in the decision-making process. Early involvement activities should be tailored to the needs of the community and be described in the communications strategy for the site. The Community Involvement Coordinator (CIC) for the site coordinates these first steps. Examples of outreach activities include giving oral presentations and poster sessions at public meetings, coordinating small group meetings or focused workshops, conducting interviews with community members to identify concerns, and disseminating fact sheets.

Ideally, the public and other stakeholders will be involved in the decision-making process long before any PRA is considered. If the community has not been previously involved, efforts should be made, in coordination with the CIC, to identify and communicate with the appropriate individuals in the community prior to the Agency's receipt of the PRA work plan. The public and stakeholders should provide input to the work plan for a PRA (Covello and Allen, 1988) (see Chapter 1).

8.2 COMMUNICATION AND PRESENTATION

8.2.1 COMMUNICATION OF PRA TO CONCERNED CITIZENS, OTHER STAKEHOLDERS, AND MANAGERS: AN OVERVIEW

When a decision is made to conduct a PRA, an important step generally is to work with citizens to plan an educational process that will help them *understand the principles of PRA* and its application in the Superfund process (see Chapter 1). This discussion may be best presented in an informal setting such as a public availability session. Because of the complex nature of quantitative uncertainty analysis and PRA, a small group meeting may be a good forum in which to discuss issues and facilitate an exchange of ideas. Such meetings provide the foundation for building trust and credibility (see Section 8.6).

In general, it is important to identify whether a Community Advisory Group (CAG) has been formed. The purpose of a Community Advisory Group is to provide a public forum for community members to present and discuss their needs and concerns related to the Superfund decision-making process. The Community Involvement Coordinator is an important member of the team and may coordinate communication plans and hand-out materials, and identify site-specific organizational issues.

A number of resources may be available to the community for help in understanding technical material. The Technical Assistance Grant (TAG) Program provides funds for qualified citizens' groups affected by a Superfund site to hire independent technical advisors to help interpret and comment on site-related information. Another program, Technical Outreach Services for communities (TOSC), uses university educational and technical resources to help Communities understand the technical issues involving the hazardous waste sites in their communities. This is a no-cost, non-advocate technical assistance program supported by the Hazardous Substance Research Centers.

The initial community meeting can serve to establish a strong rapport and facilitate the exchange of both policy and technical information needed to support a PRA. This information may include policy decisions associated with both point estimate and probabilistic approaches, as well as technical detail regarding the conceptual exposure model and the selection of distributions. This exchange may increase certainty about the assumptions made in the risk assessment. For example, the community may be able to offer insights regarding site-specific activities and sources of data not readily available to the assessor. This type of discussion allows for the free exchange of information with the public and sets the stage for future discussions. It is important that an appropriate level of detail is presented at the first meeting. Instead of overloading the audience with information, it is generally better to coordinate several meetings so that more complex policy and technical concepts can be broken down into smaller discussion topics.

The dialogue with the stakeholders may continue throughout the remedial investigation process with special attempts to involve the stakeholders in decisions concerning the PRA (see Chapter 1). Demonstrations using portable computers can be an effective tool for showing how the results of the PRA may change with changes in modeling assumptions. Such demonstrations can be included as part of the communication strategy and can be tailored to meet the community's needs.

The tiered approach for PRA presented in Chapter 1 (Figure 1-4) encourages risk assessors and risk managers to participate in discussions early in the process of developing point estimate and probabilistic approaches. If a decision is made to perform a PRA, a continuing dialogue will be useful to evaluate interim results of the PRA and determine if additional activities are warranted (e.g., data collection, further modeling). These on-going results will help assure that risk managers are aware of the details of the PRA analysis and are comfortable with the material that will be shared with the community and senior managers.

8.2.2 STEPS FOR COMMUNICATION OF THE RESULTS OF THE PROBABILISTIC RISK ANALYSIS

The results of a PRA will vary depending on the nature and extent of the assessment performed. For example, PRAs may include an analysis of variability only, uncertainty only, or both variability and uncertainty. Some analyses may involve simulations to evaluate temporal variability (e.g., Microexposure Event analysis) and spatial variability (e.g., geostatistics). The challenge for presenters is to *identify the critical information and level of detail to be presented to various audiences that may be involved in the Superfund decision making process* (i.e., senior risk managers, concerned citizens, congressional staff, responsible parties, and other users of the information) (see Exhibit 1-6).

EXHIBIT 8-1

6-STEP PROCESS FOR COMMUNICATING PRA RESULTS

1. Identify the audience
2. Identify the needs of the audience
3. Develop a communication plan
4. Practice to assure understandability
5. Post-meeting review of presentation and community feedback
6. Update information as needed for future assessments and presentations

The 6-step process described below (and summarized in Exhibit 8-1) may be repeated many times during the performance of a PRA. For communication purposes, a PRA will involve more interaction with stakeholders than a simpler point estimate risk analysis.

1. Identify the audience

The first step is to identify the audience that is interested in the information. The format and depth of the presentation of PRA information will vary according to the audience; the audience may change depending on the tier of the PRA (see Chapter 1) and the decisions being made. Accordingly, the materials communicated should be tailored to the specific needs of each of those audiences.

2. Identify the needs of the audience

The second step is to identify the information from the point estimate risk analysis and the probabilistic analysis that will address the needs of the audience (see Section 8.3). The significant information and the level of detail will, of course, vary depending on the audience (i.e., manager, assessor, peer-reviewers, responsible party, or public). This information may include a discussion of the sources of data used in the PRA, the most critical variables in the PRA identified during the sensitivity analysis, statistics and the nature of distributions, or a consideration of uncertainty (see also Section 8.5). The risk assessor should select the key information and discuss the significance of this information based on the intended audience.

3. Develop a communication plan

EXHIBIT 8-2

DEVELOPING UNDERSTANDABLE MATERIAL

- Identify main messages
- Place information in appropriate context
- Use clear formats
- Use examples and graphs
- Provide handouts and glossaries
- Present information with a minimum of jargon

The third step is to develop a plan to communicate the significant information in an easily understandable format (e.g., using graphics) (Exhibit 8-2). Adequate planning in the presentation of PRA information is essential. A thorough understanding of the design and results of the PRA will help to place the information in *proper context* and *understandable format* (Covello and Allen, 1988). Even more important than communicating the appropriate technical details, however, assessors and managers should clearly *identify the main messages* to be presented. Pilot testing a presentation with co-workers who are unfamiliar with the site can assure that the appropriate messages are being conveyed. Section 8.4 provides examples of graphics that may be useful in presentations of PRA. Handouts, glossaries, and other materials may complement a presentation and provide information for discussion following the meetings (see Section 8.4).

4. Practice to assure understandability

The fourth step is to practice the presentation to assure that the information is understandable. Staff from communication groups or public information offices within the EPA Regional Offices may help to determine whether or not the presentation addresses the needs of various audiences.

5. Post-meeting review and of presentation and community feedback

The fifth step occurs after the presentation. In addition to receiving feedback on the materials and information used in the presentation, the presenter will likely develop a sense for how well the main messages and specific technical issues were communicated.

6. Update information as appropriate for future assessments and presentations

Shortly after the meeting or briefing, modifications should be made to the materials for future presentations where appropriate. In addition, if information is obtained that is relevant to the risk assessment, this information may be included in a subsequent analysis, and the process would be repeated.

8.3 COMMUNICATING DIFFERENCES BETWEEN POINT ESTIMATE AND PROBABILISTIC RISK ASSESSMENT

One method for effectively explaining the probabilistic approach to quantifying variability and uncertainty is to employ comparisons to the more easily understood *point estimate methodology*. The communicator may focus on a specific input variable, such as drinking water intake, and first explain that with the point estimate methodology, a single value such as 2 liters per day is used to represent the entire population. However, with probabilistic analysis, those individuals consuming 1 liter per day, as well as those who consume 3 liters per day, are represented in the calculation. PRA can be used to relate such

sources of variability to the overall variability in risk. The presenter may use a concrete example of a distribution by asking the audience to identify their own water consumption level on the distribution.

When communicating results from point estimate and PRA models, an important concept to keep in mind is that both methods yield estimates of true, but unknown health risks. It is common to perceive output from quantitative models as representing the "truth", without appreciating the uncertainties in the estimates. One challenge in presenting PRA results is to determine the most effective way to communicate sources of uncertainty without undermining the trust and credibility of the assessment (see Section 8.6). Using the above example, concepts associated with uncertainty can be introduced by asking the audience if their estimate of water consumption on a specific day would be equal to their average daily consumption rate over a 1-year period. This example highlights a common source of uncertainty in exposure data (i.e., using short-term survey data to estimate long-term behavior). Section 8.5 discusses different perceptions of uncertainty.

The basic concepts of PRAs described in Chapter 1 may be used in developing presentations. Tables 1-1 and 1-2 in Chapter 1 summarize some of the advantages and disadvantages of point estimates and probabilistic approaches that should be considered when evaluating differences in the risk estimates of the two approaches. For example, point estimates of risk do not specify the proportion of the population that may experience unacceptable risks. In some cases, this proportion could be greater than 10%, or in others, less than 0.01%. In contrast, with probabilistic risk methods, statements can be made regarding both the probability of exceeding a target risk, and the level of confidence in the risk estimates. Understanding the output distribution of risks and the sensitivity analysis (Chapter 2) should help to explain the results of the assessment and address the concerns of the audience. Consider a risk management decision in which a PRG is determined from an RME risk associated with the 90th percentile of the risk distribution (i.e., the low end of the RME range). Results from the PRA used to support this decision might include extensive site characterization, site-specific data on sensitive exposure variables, and relatively narrow confidence limits on the arithmetic mean concentration.

When summarizing results of PRA, point estimates of risk generally should be presented in the same graphs and tables. It may be informative to note where on the risk distribution each of the point estimates lies. By understanding the assumptions regarding the inputs and modeling approaches used to derive point estimates and probabilistic estimates of risk, a communicator will be better prepared to explain significant differences in risk estimates that may occur. Special emphasis should be given to the model and parameter assumptions that have the most influence on the risk estimates, as determined from the sensitivity analysis.

8.4 GRAPHICAL PRESENTATION OF PRA RESULTS TO VARIOUS AUDIENCES

As the old adage, "A picture is worth a thousand words," implies, graphics can be an effective tool for presenting information. A graphic can be most easily understood by a diverse audience when it conveys a single message. In general, each graphic should be developed and modified depending on the type of presentation and audience.

- *The key is to plan how the information will be presented, select a small number of appropriate messages, and not overwhelm the audience with detail.*

Points to consider when developing graphics for public meetings, senior staff, and the press are presented below.

8.4.1 PUBLIC MEETING

For a public availability session (or meeting), care should be taken to assure that the graphics are of appropriate size and the lettering is easy to read. For example, a graphic on an 8 ½ x 11 inch sheet of paper would not be easily seen from the back of a large auditorium. It may be appropriate to present information using large posters, spaced so that the audience may move among them and discuss the posted results with the assessor or project manager. Handouts with the graphics and a glossary of terms should also be used. Avoid using slides with too much text since the information will be difficult to read and understand.

8.4.2 SENIOR STAFF

For communicating with senior risk managers, an executive summary or executive briefing may be more appropriate. This presentation should highlight major findings, compare point estimate and probabilistic results, provide sensitivity analysis results, and state the uncertainties revealed in the PRA.

A good way to present the results of a sensitivity analysis is to use a pie chart, as in Figure 8-1A. The chart shows the amount of influence a specific variable has on the final risk estimate. The figure clearly shows that the soil concentration of a contaminant has the largest effect on the risk estimate, while the soil ingestion rate has the second largest effect.

Senior level risk managers would generally be most interested in the risk level at the 50^{th}, the 90^{th}, the 95^{th}, and the 99.9^{th} percentiles (i.e., the variability in risk estimates). Managers may also wish to know the uncertainty surrounding each of the percentiles of risk. This uncertainty can be described in a table (e.g., percentiles of uncertainty in 95^{th} percentile risk) or a graphic (e.g., box-and-whisker plots). It is advisable for the risk assessor to have this information in hand. Presenting percentiles of uncertainty along with percentiles of variability can require a very busy figure or table — it is best to keep things simple. Figure 8-1B shows a cumulative probability plot for Hazard Quotient (HQ), using the box-and-whisker graphic style to show the uncertainty around selected percentiles estimates. The box shows the 25 % and 75% confidence interval for the percentile, whereas the whiskers show the 5% and 95% confidence interval. The box-and-whisker plot is simple to produce, conveys information about the skew and width of the confidence interval, and is easier to interpret for less technical audiences. In general, box-and-whisker plots are useful for summarizing results from multiple Monte Carlo simulations (e.g., results of 2-dimensional Monte Carlo analysis simulations [Appendix E] used to propagate uncertainty in parameters of probability distributions through the model).

Figure 8-1B also gives the complete CDFs of HQ associated with the 5^{th}, 50^{th}, and 95^{th} percentiles of uncertainty. This may be a useful visual aid to accompany the box-and-whisker plots. Probability density and CDFs are generally more meaningful to risk assessors and uncertainty analysts. Alternative graphics may be needed to communicate uncertainty in risk estimates based on other sources of uncertainty in the model (e.g., use of alternative probability models for exposure variables, effect of changes in the model time step, application of spatial weighting to concentration data, etc.). Additional information on communicating risks to senior EPA managers is given by Bloom et al. (1993).

8.4.3 PRESS RELEASES

For a press briefing presentation, care should be given to identify messages and develop publication quality graphics with easily understood descriptions. The risk managers generally should work with appropriate public information experts and senior managers to develop these materials and have adequate approvals before their release.

8.5 PERCEPTION OF RISK AND UNCERTAINTY

There are many individual differences in the way people regard the risks and hazards that are unavoidable in modern life. These differences have their roots in both the technical limitations of risk assessment and the idiosyncrasies of the individual human mind (Slovic, 1986). The risk assessor or risk communicator should keep in mind the general perceptions about risk held by different groups. Communications should be tailored to the audience.

The relationship between an individual's level of knowledge and perception of risk is complex. For example, an individual may know nothing about the risk associated with an activity and, on the basis of *a priori* experience, judge that risk to be negligible. Another individual may feel more threatened by his or her lack of knowledge and magnify the risk in his or her mind (Marris et al., 1997).

The presentation of uncertainty in a risk estimate can be interpreted with vastly different conclusions depending on the audience. For example, a thorough scientific account of multiple sources of uncertainty presented to a group of interested risk assessors and environmental scientists may be clearly understood. Such a group will likely conclude that the assumptions made in the risk assessment were appropriate and that the results can be used with confidence as a decision support tool.

In contrast, a similar scientific presentation given to the public may be misunderstood. Citizens are often more concerned about the impact of a risk to their situation than the uncertainty in the risk estimate. Consequently, the public is likely to react with resentment to a highly scientific presentation on uncertainty. Focusing heavily on uncertainty may cause citizens to conclude that the risk must be high. They may also conclude that the presenter is incompetent because he or she is not sure of anything, that the presenter is trying to hide something by cloaking the information in technical jargon, or that the presenter is intentionally avoiding the public's issues of concern.

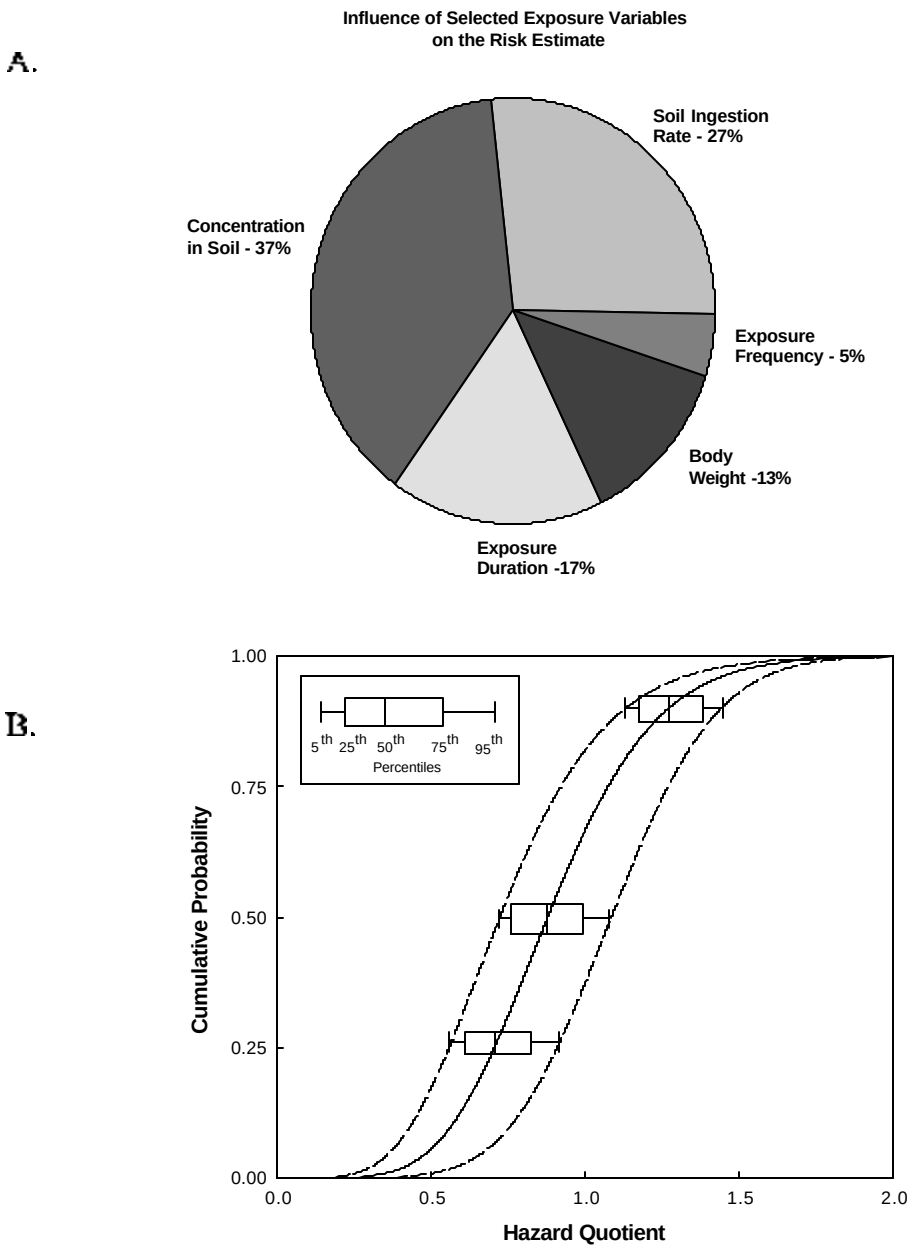


Figure 8-1. This graphic is a good way to show the results of a 2-D MCA. The top panel (A) shows a pie chart with results of a sensitivity analysis. The results represent the contribution to the variance in risk based on the square of the rank correlation normalized to 100%. Additional examples are given in Chapters 4 (Figure 4-6) and 5. The bottom panel (B) shows a method of presenting variability as a cumulative distribution function and uncertainty as box plots at the 25th, 50th, and 95th percentiles of variability. The CDFs given by the dotted lines represent the 5th and 95th percentiles of uncertainty for each percentile of variability.

A helpful presentation generally should incorporate the following steps: (1) present information about the conclusions that can be drawn from a risk assessment; (2) describe the certainty of the information that supports these conclusions; and (3) address the uncertainty and its implications for the conclusions. Graphics help people to understand uncertainty. It is extremely frustrating for decision-makers to receive detailed information on uncertainty without conclusions (Chun, 1996).

8.6 TRUST AND CREDIBILITY

The single most important quality a presenter can communicate to others is a sense of trust and credibility. Trust and credibility is based on the individual's own integrity. Building trust and credibility will serve the assessor well, whether communicating to a high-level technical audience, a risk manager/decision-maker who wishes to have the "big picture", or the lay public for whom the complexities of a risk need to be made understandable.

Credibility can only be bestowed by the public through a long history of candor and the ability to synthesize information to address the concerns and interests of an audience. If a presenter can explain an issue because he or she has insight into the topic, the presenter will gain credibility. The ability to garner trust and credibility comes from communicating with an audience at an appropriate level (Covello and Allen, 1988).

REFERENCES FOR CHAPTER 8

Bloom, D.L. et al. 1993. Communicating Risk to Senior EPA Policy Makers: A Focus Group Study. U.S. EPA Office of Air Quality Planning and Standards.

Chun, A. 1996. Strategies for Communicating Uncertainty to the Public. IBM Risk Conference Proceedings, Oct. 31, 1996.

Covello, V.T and F.W. Allen. 1988. *Seven Cardinal Rules of Risk Communication*. Office of Policy Analysis. EPA/xx/87/020.

Marris, C., I. Langford, T. Sanderson and T. O'Riordan. 1997. Exploring the "psychometric paradigm": comparisons between aggregate and individual analyses. *Risk Anal.* 17(3): 303-312.

Slovic, P. 1986. Informing and educating the public about risk. *Risk Anal.* 6(4): 403-415.

U.S. EPA. 1998. *Superfund Community Involvement Handbook and Toolkit*. U.S. EPA Office of Emergency and Remedial Response, EPA 540-R-98-007.

U.S. EPA. 1999. *Risk Assessment Guidance for Superfund: Volume 1 - Human Health Evaluation Manual*. Supplement to Part A: Community Involvement in Superfund Risk Assessments. EPA/540/R-98/042, March.